Remote sensing of freshwater phytoplankton

Fergus Charlton

Submitted for the degree of PhD

University of Edinburgh

1998

Declaration

This thesis, unless otherwise stated, is solely the work of the undersigned.

Fergus Charlton

28-Oct-1998

Acomb

Northumberland

Abstract

The quality of inland water bodies can become adversely affected by the growth of high concentrations of phytoplankton. Traditionally water managers have relied upon spot sampling to assess the potential for bloom development, performing measurements of both phytoplankton concentration and composition. Phytoplankton concentration indicates the stage of bloom development. Phytoplankton composition allows assessment of the particular water quality problems associated with blooms dominated by different phytoplanktonic classes. The synoptic nature of broad-band remote sensing has provided water managers with a tool to map lake-wide variations in phytoplankton concentration in near real time. The development of hyperspectral airborne remote sensors has further increased the scope for information retrieval. This has given remote sensing the potential to map spatial variations in phytoplankton composition which would be a significant advantage for water managers.

This study researches the potential for using hyperspectral remote sensing to identify the phytoplanktonic composition of a freshwater bloom. Six novel analytical techniques were developed to identify phytoplankton class from reflectance spectra. These techniques offer the water manager a variety of means to identify the dominant phytoplankton class in a target water body.

Identification of phytoplankton class is possible because certain photosynthetic pigments contained within phytoplankton cells are taxonomically significant, being indicative of a particular class. The detection of these pigments can be used to identify the presence of a particular phytoplanktonic class in an aquatic system. It is possible to identify these pigments using optical methods because they exhibit unique spectral absorption signatures. Such pigment absorption features are manifest in the composite reflectance signature from water bodies as measured by remote sensing instruments. However, due to the presence of the spectral features from other photosynthetic pigments and the other optically active components of water bodies, extracting from reflectance spectra the spectral information pertaining to individual class marker pigments can be difficult. The phytoplankton class identification techniques presented in this study were developed using absorption and reflectance spectra from pure cultures of phytoplankton. The reflectance spectra were measured in the controlled environment of a laboratory based experimental tank designed for this study. The class identification techniques were tested on field and airborne reflectance spectra measured from a eutrophic inland lake.

Acknowledgements

I acknowledge the support from Natural Environment Research Council under Studentship GT4/93/26/a.

I would like to thank: my supervisors Tim and Glen; all the staff at Ferry House especially Julie, Eric, Mitzi, Diane, Christine, Peter, Peter, Colin, Paul, Paul, Mike and Martin; Irene from the NRA; Jerry from Sutton Bonnington; Martin from Wolverhampton; Jim and Mike from Monkswood; Arnold and Erin from Amsterdam; Herman from Nieuwersluis.

Thanks to my mother and father for providing financial support (and a home) during this study. Both were much appreciated.

This work is dedicated to Emma and Poppy.

List of abbreviations, terms and symbols

General abbreviations

AOV	=	area of view
Case I waters	=	Waters where phytoplankton are the dominant optically active parameter
Case II waters	=	Waters where phytoplankton are not the dominant optically active parameter
DM	=	Diatom Media (phytoplankton culture media)
DOC	=	dissolved organic carbon
DYS	=	dissolved yellow substances
F	=	F-ratio value
FOV	=	field of view
HPLC	=	high performance liquid chromatography or high pressure liquid chromatography
IR	=	infra-red wavelengths
JM	=	Jaworski's Media (phytoplankton culture media)
nm	=	nanometer (10^{-9} m)
PAR	=	photosynthetically active radiation (400 to 700 nm)
UV	=	ultra-violet (wavelengths between 200 and 400 nm)

Optical and radiometric properties

ϕ	=	azimuthal angle
ε	=	factor pertaining to spectrophotometer sample chamber which determines the amount of scattered light received by the photomultiplier
ρ	=	size parameter
λ	=	wavelength (nm)
θ	=	zenith angle
$\chi\Delta(\lambda)$	=	relative scattering
$\Delta\chi(\lambda)$	=	relative scattering
$\chi(\lambda)$	=	apparent absorption
$\beta(\theta)$	=	volume scattering function
ε^*	=	maximised ε
$a^*(\lambda)$	=	specific absorption spectra ($m^2\ \mu g^{-1}$)
$a_{CM}(\lambda)$	=	absorption by cellular material (m^{-1})
$a_{DYS}(\lambda)$	=	absorption by dissolved yellow substances (m^{-1})
$a_{H2O}(\lambda)$	=	absorption by water (m^{-1})
$a_{PHY}(\lambda)$	=	absorption by phytoplankton (m^{-1})
$a_{SES}(\lambda)$	=	absorption by seston (m^{-1})
$a_{SOL}(\lambda)$	=	absorption by an equivalent solution (m^{-1})
$a_{TOT}(\lambda)$	=	total absorption (m^{-1})
$a_{TRI}(\lambda)$	=	absorption by tripton (m^{-1})
$b(\lambda)$	=	scattering (m^{-1})
$b_b(\lambda)$	=	backscattering (m^{-1})
$b_{bPHY}(\lambda)$	=	backward scattering by phytoplankton (m^{-1})
$b_{fPHY}(\lambda)$	=	forward scattering by phytoplankton (m^{-1})
$b_{PHY}(\lambda)$	=	scattering by phytoplankton (m^{-1})
$b_{TOT}(\lambda)$	=	total scattering (m^{-1})
$c(\lambda)$	=	beam attenuation (m^{-1})

$c_{PHY}(\lambda)$	=	beam attenuation by phytoplankton (m^{-1})
$D(\lambda)$	=	absorbance (dimensionless)
$Ed_a(\lambda)$	=	above surface downwelling irradiance (Wm^{-2})
$Ed_{aDIF}(\lambda)$	=	the diffuse fraction of above surface downwelling irradiance (Wm^{-2})
$Ed_{aSUN}(\lambda)$	=	the direct fraction of above surface downwelling irradiance (Wm^{-2})
$Eu_w(\lambda)$	=	upwelling below surface irradiance (Wm^{-2})
$F(\lambda)$	=	the fraction of $Ed_a(\lambda)$ that is diffuse (= $Ed_{aDIF}(\lambda)$ / $Ed_{aSUN}(\lambda)$)
g_{440}, g_{340}	=	a_{DYS}(440 nm) or a_{DYS}(340 nm)
$Kd(\lambda)$	=	downwelling attenuation coefficient (m^{-1})
Kd(PAR)	=	downwelling attenuation coefficient integrated across PAR wavelengths (m^{-1})
$Lo(\lambda)$	=	radiance reflected from Lambertian reference panel ($Wm^{-2}sr^{-1}$)
$Lr_a(\lambda)$	=	surface reflected radiance ($Wm^{-2}sr^{-1}$)
$Lu_a(\lambda)$	=	upwelling surface-leaving radiance ($Wm^{-2}sr^{-1}$)
$Lu_w(\lambda)$	=	upwelling below surface radiance ($Wm^{-2}sr^{-1}$)
$Lz_a(\lambda)$	=	upwelling radiance measured by radiometer viewing in the nadir ($Wm^{-2}sr^{-1}$) (= $Lra(\lambda)$ + $Lu_a(\lambda)$)
Q	=	the ratio of irradiance to radiance (sr)
r	=	pathlength (m^{-1})
$R(\lambda)$	=	reflectance
$R(\lambda)$(+0.01m)	=	above surface irradiance reflectance at +0.01 m
$R(\lambda)$(-0.01m)	=	below surface irradiance reflectance at -0.01 m
$R_{PHY}(\lambda)$	=	reflectance by phytoplankton
z	=	depth (m)
Z_{SD}	=	Secchi disk depth (m)

Phytoplankton species abbreviations

Ana	=	*Anabaena*
Ast	=	*Asterionella*
Chl	=	*Chlorella*
Fra	=	*Fragilaria*
Mic	=	*Microcystis*
Sel	=	*Selenastrum*

Phytoplankton pigment abbreviations

C abbreviates association with Cyanophytes, R abbreviates association with Rhodophytes and B abbreviates association with Bacillariophytes.

ADX	=	Anhydrodiatoxanthin
ALX	=	Alloxanthin
ANX	=	Antheraxanthin
APC	=	Allophycocyanin or C-APC
APN	=	Aphanin
ATX	=	Astaxanthin
AZP	=	Aphanizophyll
B-CART	=	β-Carotene
CCX	=	Crocoxanthin
CHLa	=	Chlorophyll-*a*
CHLb	=	Chlorophyll-*b*
CHLc	=	Chlorophyll-*c*, -c_1, -c_2, and -c_3
CHLd	=	Chlorophyll-*d*
CHLMgDVP	=	Chlorophyll MgDVP
CLX	=	Calloxanthin
CPX	=	Cryptoxanthin

CTX	=	Canthaxanthin
DDX	=	Diadinoxanthin
DNX	=	Dinoxanthin
DTX	=	Diatoxanthin
ECN	=	Echineone
FCX	=	Fucoxanthin
G-CART	=	γ-Carotene
HTX	=	Heteroxanthin
LUT	=	Lutein
MDX	=	Monadoxanthin
MTC	=	Mutatochrome
MXX	=	Myxoxanthin
MYX	=	Myxoxanthophyll
NOX	=	Neoxanthin
NPD	=	Neoperidinin
OCX	=	Oscillaxanthin
PCC	=	Phycocyanin and C-PCC, R-PCC and B-PCC
PCE	=	Phycoerythrin and R-PCE and B-PCE
PEC	=	Phycoerythrocyanin
PRD	=	Peridinin
PRX	=	Pyroxanthin
PSX	=	Prasinoxanthin
SIP	=	Siponein
SPX	=	Siphonaxanthin
TRX	=	Taraxanthin
VCX	=	Vaucheriaxanthin
VLX	=	Violaxanthin
ZEX	=	Zeaxanthin

Contents

Chapter One Introduction

In the presence of sufficient nutrients, and given the right innoculum and growing conditions, a water body can become adversely affected by large populations of phytoplankton. Large concentrations (or blooms) of phytoplankton have poor aesthetic qualities and can severely alter the balance of dissolved oxygen over the diurnal period through oxygen production and oxygen uptake. Furthermore, should the water be required as a potable source, then certain species of phytoplankton can block filters, impart severe odour and taste problems to the water, and in some instances be toxic to higher forms of life (Tebbutt 1983). The problems that a bloom imposes upon water managers are therefore related not just to the concentration of the bloom but also upon the type of phytoplankton present in that bloom. With rising interest in the importance of environmental issues, water managers have become increasingly concerned about the condition of inland water bodies with a focus on the problem of eutrophication (Van Stokkom *et al.* 1993, NRA 1993). The cost to water managers of treating phytoplankton blooms is significant.

Both research workers and water managers have used measurements of water colour when assessing water quality (Bukata *et al.*1985). Airborne remote sensing of freshwaters is an extension of this colour measurement technique. The nature of remote sensing is such that it offers a means to circumvent many of the cost and time problems associated with traditional spot sample monitoring. Moreover, remote sensing is unique in that it provides a synoptic coverage of water bodies and has the potential for near real-time monitoring that will lead to improved understanding of aquatic systems and is necessary for the effective amelioration of water quality problems associated with phytoplankton blooms (Millie *et al.* 1992).

The presence of optically active water quality components can be detected and quantified in the signal measured by remote sensors. Airborne remote sensors simultaneously view broad swaths of a lake's surface thereby allowing the mapping of variations in these parameters across that surface (Dekker *et al.* 1992a, Gitelson *et al.* 1994). Of the various water quality parameters of interest to water managers remote sensing has proved to be a particularly effective means of monitoring variations in phytoplankton concentration (Millie *et al.* 1992), therefore remote sensing offers the ability to monitor the processes of eutrophication in inland waters (Gitelson *et al.* 1993). Remote sensing has also been shown to be beneficial in the optimisation of the more traditional spot sampling regimes (Dekker *et al.* 1992a), this is especially true when lakes have a tendency to exhibit spatial patterns of variable phytoplankton concentrations (George 1993).

Most freshwater phytoplankton have been taxonomically classified into: classes, those groups of phytoplankton *genera* exhibiting broadly similar characteristics, in terms of morphology and pigmentation (*e.g.* the Chlorophytes or Greens); *genera* which are groups of phytoplankton species

exhibiting even more similarity (*e.g.* the *genus Selenastrum*); and finally the single species (Bold and Wynne 1985). For example the species *Selenastrum capricornitum* is a member of the *genera Selenastrum* which is a member of the class Chlorophyceae. This classification is relevant because, not only does the concentration of freshwater phytoplankton in a lake change over a year, but so does the phytoplankton composition and phytoplankton classification is required to describe that composition.

Reynolds (1980, 1988) develops the idea of serial succession of freshwater phytoplankton populations using the concept of functional groups. Functional groups of freshwater phytoplankton are combined on the basis of their ability to exploit different environmentally induced 'niches' such as the ability to grow when light or phosphate is limiting. These functional groups tend to fall within the taxonomically defined phytoplankton classes. The prevailing environmental conditions during a year are responsible for determining which of the functional groups is in ascendance and which is in decline. Figure 1 summarises changes likely to occur in temperate lakes of differing trophic status as environmental conditions change to the benefit of specific functional groups. A simplified version of this successional change would be: in the spring, when the lake is not stratified, it is likely to be dominated by Diatoms which tend to have morphologies which allow them to harvest the weaker energy of the lower angled sun more efficiently; in late spring Chlorophytes begin to be dominant as sun strength increases and the Diatoms either sink out the water column following the onset of stratification or become restricted by silica availability; as epilimnon water temperatures increase during late summer Cyanophytes dominate successfully competing for all the available phosphorous; the break-down of stratification during autumn wind events causes falling temperatures and the decline of the Cyanophytes.

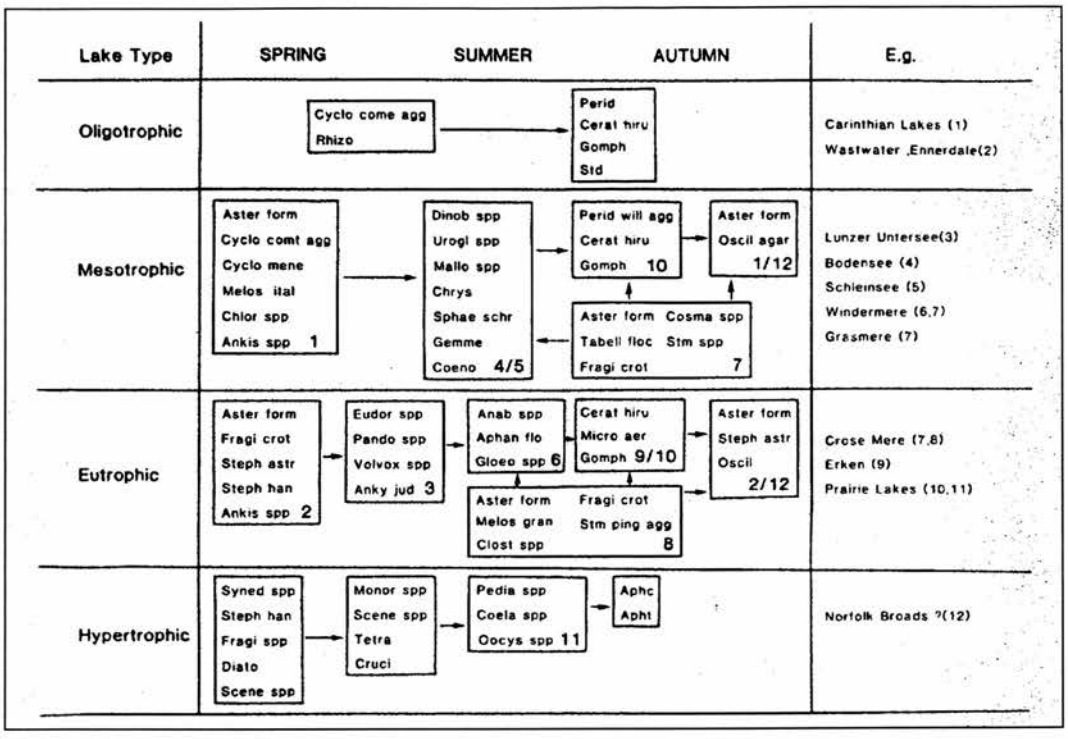

Figure 1. (Reproduced with permission from Reynolds 1982). The generalised successional pathways in lakes of different trophic status. Each box represents a functional group of phytoplankton. The boxes represent the functional groups of species one or more of which may be abundant. The bold number refer to the functional groups of Reynolds (1980). Abbreviations to species: Ankist = *Ankistrodesmus* spp; Anky = *Ankyra judayi*; Anab = *Anabaena* spp; Aphan flo = *Aphanizomenon flos-aquae*; Aphc = *Aphanocapsa*; Apht = *Aphanothece*; Aster from = *Asterionella formosa*; Cera hir = *Ceratium hirundinella*; Chlo = *Chlorella* spp; Chrys = *Chrysochromulina*; Clost spp = *Clostrium* spp; Coela = *Coelastrum* spp; Coeno = *Coenococcus*; Cosma = *Cosmarium* spp; Cruci = *Crucigenea* spp; Cyclo come agg = *Cyclotella comensis* aggregate; Cyclo comt agg = *Cyclotella comta* aggregate; Cyclo mene = *Cyclotella meneghiniana*; Diato = *Diatoma*; Dinob = *Dinobryon*; Eudor spp = *Eudorina spp*; Frago spp. = *Fragilaria* spp; Gemme = *Gemellicystis neglecta*; Gloe = *Gleotrichia*; Gomph = *Gomphosphaeria*; Mallo = *Mallomonas*; Melos gran = *Melosira granulata*; Melos ital = *Melosira italica*; Micro aer = *Microcystis aeruginosa*; Monor spp. = *Monoraphidium* spp; Oocys spp = *Oocystis* spp; Oscil = *Oscillatoria aagardhii*; Pando spp = *Pandorina* spp; Pedia spp = *Pediatrum* spp; Rhiz = *Rhizoslenia*; Scene = *Scenedesmus* spp; Spae schr = *Spaerocytis schroeteri*; Std = *Staurodesmus*; Steph astr = *Stephanodiscus*; Steph han = *Stephanodisicus*; Stm spp = *Staurastrum* spp; Syned spp. = *Synedra* spp; Tabell floc = *Tabellaria flocculosa*; Tetra = *Tetrastrum*; Urogl spp= *Uroglena* spp; Volvox spp = *Volvox* spp.

Lake phytoplankton populations are therefore dynamic in terms of both their concentration and their composition. This quantitative and qualitative variability in the phytoplankton population has implications for both those involved in the management of freshwater bodies and remote sensors. It would therefore be useful if remote sensing could also provided information pertaining to the type of

phytoplankton in a bloom. Indeed it is likely that the accuracy with which phytoplankton biomass concentration is measured by remote sensing will be improved if consideration is given to the type of dominant phytoplankton in the target bloom. Knowledge of the type of phytoplankton would therefore allow the water manager to apply the optimal algorithm for monitoring individual blooms using remote sensing techniques.

It can be concluded that, for remote sensing to be a truly effective tool for the water manager it must be able to provide information not only on the spatial variation in phytoplankton concentration, but also the temporal variation in phytoplankton type.

Remote sensing has the potential to identify phytoplankton type due to the colour differences induced by the presence of various photosynthetically active pigments. These pigments have discreet absorption spectra, which remove visible light from the underwater light climate thereby altering the spectral quality of the water leaving upwelling radiance flux. Remote sensors measure the spectral quality of this flux (Dekker 1993, Kirk 1994). The optical properties of phytoplankton vary widely in both magnitude and spectral quantity according to variations in their pigment content. The identification of phytoplankton on the basis of their colour depends upon the ability to extract information on their pigment composition from measurements of that colour, which must be combined with a knowledge of pigment class associations (Johnsen *et al.* 1994). Fortuitously, phytoplankton *genera* from the same class have similar pigmentation, it should therefore be possible to use remote sensing to identify phytoplankton class on the basis of the identification of class marker pigments (Bricaud and Morel 1986, Haardt and Maske 1987, Bidigare *et al.* 1989, Hoepffner and Sathyendranath 1991, Hoepffner and Sathyendranath 1993). Remote sensing could therefore be used to provide maps of the spatial variation in phytoplankton biomass along with variations in type across the surface of the target water body. This would be of more use to water managers tackling blooms than the traditional spot sample.

Very few studies of the potential for remote sensing to monitor phytoplankton type have been attempted. Brown *et al.* (1995) successfully predicted oceanic phytoplankton composition through remote sensing by calculating the ratio of euphotic depth to mixed-layer depth (both of which can be ascertained through the use of traditional satellite remote sensing techniques) and then using this ratio to predict the relative abundance of three major oceanic phytoplankton classes. Although this was an indirect technique, it was successful because the various classes had different preferences as to their preferred environmental conditions and these conditions were prevailing at distinct depths in the stratified ocean under study. To assess the spatial extent, biomass concentration, and class type, in inland aquaculture impoundments, Millie *et al.* (1992) used airborne remote sensing to monitor pigment composition. Their approach to class identification was based on an algorithm for the detection of a Cyanophyte marker pigment. The poor performance of this algorithm severely limited their ability to record variations in phytoplankton class. Dekker (1993) also used a hyperspectral

remote sensing instrument to develop and apply a Cyanophyte marker pigment retrieval algorithm to detect the presence or absence of Cyanophytes in Dutch inland waters.

The potential for remote sensing to identify phytoplankton type has recently been improved with the development of airborne hyperspectral remote sensing instruments (Zacharias *et al.* 1992). These instruments sample the visible spectrum in many discrete narrow bands as opposed to the more traditional remote sensing instruments which were restricted to sampling the same spectrum in a few broad bands.

1.1 Research aims and objectives

The aim of this research was to further develop remote sensing as a tool for water managers tackling the problem of freshwater phytoplankton blooms. The were two objectives to this aim. The first was the development of robust remote sensing techniques capable of distinguishing between the different classes of freshwater phytoplankton that are applicable to the current range of airborne remote sensors. These new techniques were to take advantage of the information rich data from hyperspectral remote sensing instruments. They should also consider low spectral resolution data so that they could be applied to the range of airborne remote sensing instruments currently available. This objective was to be achieved by the measurement of the absorption spectra for pure cultures of freshwater phytoplankton from which those absorption features relating to the presence of class marker pigments could be highlighted. If the presence of these pigment features in the reflectance spectra from pure phytoplankton cultures measured under controlled laboratory conditions could be detected then they could be used to indicate the presence of a particular class.

The fundamental approach taken to achieve this objective was based on the assumption that phytoplankton pigment composition determines the phytoplankton absorption spectra which in turn determines the phytoplankton reflectance spectra. The corollary of this argument is that it should be possible to identify phytoplankton type from the pigment signature of a phytoplankton bloom as recorded in its reflectance spectra.

A second objective was to ascertain how various remote sensing algorithms for the quantification of a phytoplankton bloom were affected by qualitative variations in the composition of that bloom. Supplementary to this objective was a comparison of the empirical and universal CHL*a* retrieval algorithms.

Chapter Two Theory

A brief theoretical introduction is necessary to understand the problems associated with the identification of phytoplankton class from reflectance spectra and to follow the methods developed within this study to tackle these problems. This chapter introduces the fundamental concepts behind remote sensing of aquatic targets. A description of how the optical properties of the various components of a water body are responsible for the spectral variation of the water's reflectance spectra is given. This variation is the key to quantitative and qualitative water quality assessment through remote sensing. The strong inverse relationship between phytoplankton absorption and phytoplankton reflectance is highlighted. This relationship is used through-out this study as class identification routines are first developed on the easy to measure phytoplankton absorption spectra then applied to reflectance spectra. The chapter ends with a discussion of the problems of remote sensing in the field, describing how experimental laboratory based remote sensing is more appropriate for the development of phytoplankton class identification routines than field measurements of phytoplankton reflectance.

2.1 The inherent optical properties of natural water bodies

On penetrating a water body a photon from the sun will either be scattered back out of the water or absorbed by the various constituents of the water. The processes of absorption $a(\lambda)(m^{-1})$, and scattering $b(\lambda)(m^{-1})$, along with their combination, attenuation $c(\lambda)(m^{-1})$, and the volume scattering function $\beta(\theta)(m^{-1}sr^{-1})$ (which describes the likely angle at which a photon may be scattered) are known as the inherent optical properties of a water body (Preisendorfer 1986, Bricaud and Morel 1986). These four properties of the underwater light climate are termed inherent optical properties because the magnitudes of these coefficients are independent of the illuminating light field (Kirk 1994), and as such are the fundamental variables controlling the spectral quality of the underwater light climate. Understanding these inherent optical properties is essential for the interpretation of the apparent optical properties (§ 2.2) which include water colour and reflectance (Hoepffner and Sathyendranath 1993).

In any water body there are various optically active components which have their own inherent optical properties and are therefore responsible for the spectral quality of the underwater light climate. These optically active components can be grouped into: water molecules; dissolved organic substances; particulate organic substances including phytoplankton and decaying phytoplankton; and inorganic particulate substances.

A particular characteristic of inherent optical properties is that they are strictly additive (Morel and Bricaud 1986) hence they all combine to form the total absorption $a_{TOT}(\lambda)$ and scattering $b_{TOT}(\lambda)$ exhibited by a water body. Consequently, when remote sensing a body of water each optically active

component requires consideration because of its influence on the underwater light climate and therefore on reflectance.

2.1.1 The absorption process

When a photon passes within the vicinity of a molecule there is a finite possibility that it will be absorbed. Absorption causes the energy level of the molecule to be raised from its ground state by an amount corresponding to the energy of the photon, the molecule will then dissipate that energy returning to its ground state (Kirk 1994). Increments and decrements of energy occur at three levels: rotational; vibrational; and electronic, with various combinations of these energy states being conceivable. These combined and often overlapping energy levels are responsible for the Gaussian shaped absorption spectrum of individual molecules. Different molecular structures have different series of energy levels consequently their absorption spectra have a distinctive wavelength dependency.

In the absence of multiple scattering, absorption should conform with Beer's Law, that is the absorption by a suspension of n particles is n times that of a 'suspension' of one particle (Duysens 1956). Absorption spectra for the various optically active components found in water bodies have been measured on many occasions using spectrophotometers (Yentsch 1962, Davis-Colley *et al.* 1986, Malthus and Dekker 1990).

2.1.2 The scattering process

A photon that has interacted with a molecule but has not been absorbed will be scattered *i.e.* it will experience a change in trajectory. The scattering process can be split into two processes: scattering by particles and scattering by density fluctuation. Particle scattering results from either refraction, reflection or diffraction, or a combination of these (Kirk and Tyler 1986). This form of scattering is relatively insensitive to wavelength. Particle scattering at small forward angles can be attributed to diffraction, while scattering at larger angles is attributed to external reflection or transmission with refraction. In most natural water bodies scattering is dominated by particles greater than 2 μm in diameter which results in predominantly forward scattering (Kirk 1994).

Density fluctuation scattering is the result of a photon passing through media of two different densities hence of two different refractive indices leading to refraction. This form of scattering is often termed Mie scattering after Mie (1908) and simplified by Van de Hulst (1956). Mie scattering is dependent on photon wavelength in relation to particle size, and on the particle shape and refractive index. Density fluctuation scattering is absorption dependent according to the anomalous dispersion theory of Kettler-Helmholtz which shows that within an absorption band scattering is reduced because of the effect that the absorption has on the imaginary part of the refractive index (Morel and Bricaud 1986). Hence, because absorption is wavelength dependent, this form of scattering will also be wavelength dependent. The scattering properties of various optically active components have been most

successfully measured using spectrophotometers fitted with integrating sphere attachments (Davis-Colley *et al.* 1986, Dekker 1993).

The probability of scattering occurring at a particular angle is described by the volume scattering function $\beta(\theta)$ where θ is the zenith angle (because the volume scattering function is the same in both the zenith and azimuthal planes only the zenith angle is necessary to fully describe $\beta(\theta)$). The modelling work of Kirk (1981a, 1981b) has shown that scattering in the backward direction is the most important scattering direction for the generation of the spectral signature (reflectance). Integrating $\beta(\theta)$ from 0 to 0.5π gives the backscattering coefficient b_b:

$$b_b = 2\pi\int_{\pi/2}^{\pi}\beta(\theta)\sin\theta\delta\theta \quad \text{.....2.1.2.1}$$

and integrating from 0.5π to π gives forward scattering (Morel and Bricaud 1986). The $\beta(\theta)$ for the different optically active components differ. As a general rule the smaller the refractive index the more pronounced the general asymmetry of the $\beta(\theta)$ (Morel and Bricaud 1986).

2.1.3 The inherent optical properties of water molecules

The inherent optical properties of water molecules will always influence the apparent optical properties of any body of water. Water molecules have a distinctive absorption spectrum $a_{H2O}(\lambda)$, which has been measured by Smith and Baker (1981) and more recently by Hakvoort (1994) who found that there was a temperature dependency in $a_{H2O}(\lambda)$. The absorption of pure water $a_{H2O}(\lambda)$ is low in the blue and green wavelengths, becoming increasingly significant beyond 550 nm with shoulders at 600 and 660 nm and a significant peak rising from 700 to 750 nm. All the features in the $a_{H2O}(\lambda)$ can be attributed to harmonics in the vibration of the molecules O - H bonds. The strong near-IR $a_{H2O}(\lambda)$ places an upper limit on the wavelengths (approximately at 720 nm) that can be realistically used for aquatic remote sensing of water quality (with the exception of monitoring temperature and high concentrations of tripton in the IR).

Scattering by water $b_{H2O}(\lambda)$ is density fluctuation scattering which can be theoretically shown to increase inversely with the 4^{th} power of the wavelength (Hakvoort 1994). $b_{H2O}(\lambda)$ imparts a turquoise blue colour to the purest natural waters. Density fluctuation induced scatter should occur to the same extent in the forward and backward directions, and, given that suspended particles scatter predominantly in the forward direction, $b_{H2O}(\lambda)$ can be responsible for a significant part of the upwelling flux (Kirk and Tyler 1986). $b_{H2O}(\lambda)$ is magnitudinally lower than that of $a_{H2O}(\lambda)$.

2.1.4 The inherent optical properties of dissolved organic substances

Dissolved organic substances arise from decaying organic matter. They can be produced autochthonously from the decay of a phytoplankton crop or they can be of allochthonous origin consisting of humic or fulvic acids washed in from the surrounding catchment (Tipping *et al.* 1988).

Dissolved organic substances are referred to by various names including dissolved yellow substances (DYS), coloured dissolved organic carbon (DOC), gilvin or gelbstoff (Dekker 1993).

Absorption by DYS $a_{DYS}(\lambda)$ decreases exponentially as a function of wavelength being significant in the blue and near-UV wavelengths but very low or absent in the red. Consequently the presence of coloured DOC imparts a yellow-brown colour to the water (Davis-Colley and Vant 1987, Carder *et al.* 1989). $a_{DYS}(\lambda)$ is important because: firstly, it is in direct competition with aquatic life for light; and secondly, $a_{DYS}(\lambda)$ imparts a yellow-brown hue to the water which can pose an aesthetic problem to water managers; and finally, the presence of significant amounts of coloured DOC can preclude the use of the blue wavelengths when developing algorithms for the purpose of retrieving information on water quality when remote sensing inland waters (Dekker 1993).

$a_{DYS}(340)$ and $a_{DYS}(440)$ (defined by Tipping *et al.* (1988) as g_{340} and by Kirk (1994) as g_{440} respectively) have both been used as comparative indices of the relative importance of absorption by DOC for different water bodies (Davis-Colley and Vant 1987, Kirk 1994). The exponential form of $a_{DYS}(\lambda)$ can be described according to:

$$a_{DYS}(\lambda) = a_{DYS}(R) \exp^{(S(R - \lambda))} \quad \text{.....2.1.4.1}$$

where R is a reference wavelength and S is the slope parameter characterising the regression of the natural logarithm of $a_{DYS}(\lambda)$ against λ (Bricaud *et al.* 1981, Davis-Colley and Vant 1987). Because plots of $\ln(a_{DYS})(\lambda)$ are linear across most of the visible spectrum (Davis-Colley and Vant 1987, Dekker 1993) Equation 2.1.4.1 allows the full spectrum of $a_{DYS}(\lambda)$ to be described merely by the S and R parameters.

There is no published information on the scattering by DYS $b_{DYS}(\lambda)$, but it will most probably take a similar form to $b_{H2O}(\lambda)$ because such molecular scattering will be in the realm of density fluctuation scattering.

2.1.5 The inherent optical properties of seston

Seston is the term used to cover all suspended particles in a water body. It is typically sub-divided into tripton (principally inorganic particles but including decaying organic matter) and phytoplankton (Kirk 1994). The presence or absence of one or other of these sub-divisions has important consequences for the optical properties of a water body. When only phytoplankton are present their optical properties tend to dominate the underwater light climate. Such waters are referred to as Case I waters (Morel and Prieur 1977). Case II waters are not dominated by phytoplankton so the optical properties of such waters are spectrally more complex being a combination of the optical properties of tripton and phytoplankton (Morel and Prieur 1977, Lee *et al.* 1994).

2.1.5.1 The inherent optical properties of tripton

Tripton particles enter freshwater bodies from atmospheric deposition and from catchment runoff, they can also be resuspended from the lake bed by wave action. Tripton absorption $a_{TRI}(\lambda)$ is dependent on the mineral composition of the particles and on whether the particles are coated in organic matter (Kirk 1994). Scattering by large tripton particles $b_{TRI}(\lambda)$ is particle scattering (§ 2.1.2) so has no wavelength dependency, however, smaller tripton particles will scatter the shorter wavelengths preferentially.

2.1.5.2 The inherent optical properties of phytoplankton

Phytoplankton are uni-cellular or multi-cellular plants which live in a variety of environments including freshwater lakes. As mentioned in the Introduction chapter phytoplankton are sub-divided into classes, *genera* and species. Traditionally, taxonomists have classified phytoplankton using three methods of identification. Microscopic identification is the most frequently used method with phytoplankton cells being viewed under magnification and identified to *genera* and species level on the basis of their morphology (Lund 1959, Youngman 1971). This gives a thorough, quantitative assessment of the phytoplankton crop, however it is skill intensive requiring the operator to have knowledge of a wide variety of different cell morphologies. Flow cytometry has also been used to classify phytoplankton on the basis of their fluorescence signal (Butterwick *et al.* 1982). More recently the photosynthetic pigments of the phytoplankton have been extracted and analysed by high pressure liquid chromatography (HPLC) (Mantoura and Llewelyn 1983, Gieskes and Kraay 1983, Wilhelm *et al.* 1991). HPLC identifies photosynthetic pigment composition from which phytoplankton *genera* can be identified. An alternative, less traditional approach to phytoplankton identification was employed by Johnsen *et al.* (1994) who used discriminant analysis to distinguish classes from phytoplankton absorption spectra.

Phytoplankton absorb photons to provide the energy for photosynthesis. *In vivo* phytoplankton absorption $a_{PHY}(\lambda)$ is a function of the size and shape of the phytoplankton cell structure and of the cell's photosynthetic pigment composition (Morel and Bricaud 1986). Although the cellular material of the phytoplankton will have some absorption properties they are unlikely to be as significant as the absorption by pigments $a_{PIG}(\lambda)$. Consequently pigment composition is often the principal factor controlling variation in $a_{PHY}(\lambda)$ (Roesler *et al.* 1989, Hoepffner and Sathyendranath 1991). On a first approximation $a_{PHY}(\lambda)$ is the sum of the absorption by the individual pigments $a_{PIG}(\lambda)$ according to:

$$a_{PHY}(\lambda) = \sum_{i=1}^{n} C_i a_i'(\lambda) \quad \text{.....2.1.5.2.1}$$

where C_i and $a_i'(\lambda)$ are the concentration and the specific absorption coefficient of the i^{th} pigment respectively. The spectral form of $a_{PHY}(\lambda)$ is therefore a function of pigment composition (which determines the wavelength location of absorption features) and pigment concentration (which determines the magnitude of these absorption features).

Variations in cell size and shape (*i.e.* cell morphology) can be significant both on a species and on a class level, with different species from the same class exhibiting considerably different cell morphologies. Variations in pigment composition tend to be more significant at the class level, with considerable within-class similarity in pigment composition (Prieur and Sathyendranath 1981, Prezelin and Boczar 1986, Rowan 1989). One consequence of this intra-class pigment similarity is that species from the same class are similarly coloured because their particular pigment compositions have similar light absorbing properties. The corollary of this is that much of the spectral variation in the location of $a_{PHY}(\lambda)$ spectral features is a function of phytoplankton class. Furthermore the $a_{PHY}(\lambda)$ for any species could be reconstructed if the pigment composition and concentration, and the individual *in vivo* $a_{PIG}(\lambda)$ were known. However little is known about the *in vivo* $a_{PIG}(\lambda)$ (Hoepffner and Sathyendranath 1991) and the $a_{PIG}(\lambda)$ extracted in solvents do not match *in vivo* $a_{PIG}(\lambda)$ due to hypsochromatic shifts (Johnsen *et al.* 1994). One final consideration is that estimates of $a_{PIG}(\lambda)$ from a single culture grown in the laboratory may not be representative of the $a_{PIG}(\lambda)$ for natural populations due to the different conditions imposed by the two growing environments.

Phytoplankton pigments absorb throughout the visible spectrum, however due to both variations in pigment concentration and pigment absorption efficiencies, $a_{PHY}(\lambda)$ have very distinct spectral forms. Typically $a_{PHY}(\lambda)$ is strong in the blue and far-red but weak at green wavelengths (hence many phytoplankton appear green). There are no photosynthetically active pigments absorbing beyond 750 nm (Davis-Colley *et al.* 1986). The overlapping nature of the various pigment absorption spectra that combine to form $a_{PHY}(\lambda)$ can mean that the contributions of individual $a_{PIG}(\lambda)$ to $a_{PHY}(\lambda)$ are masked (Hoepffner and Sathyendranath 1991).

Many techniques have been developed for the measurement of *in vivo* $a_{PHY}(\lambda)$. There are two main problems associated with the measurement of *in vivo* $a_{PHY}(\lambda)$: firstly, unless the sample is from a laboratory culture, phytoplankton tend not to exist in isolation from the tripton fraction of suspensoids, therefore to obtain $a_{PHY}(\lambda)$ from $a_{SES}(\lambda)$ it is necessary to correct for $a_{TRI}(\lambda)$; secondly, because phytoplankton are particles and can not be placed into solution without modifying their absorption properties, a problem arises because of particulate scattering within the sample chamber of the spectrophotometer which is anomalously recorded by the spectrophotometer as absorption, thus over estimating true *in vivo* $a_{PHY}(\lambda)$. A further problem arises from low concentrations of phytoplankton present in the natural environment which can lead to weak absorption signals and hence difficulties with the signal-to-noise ratio of most spectrophotometers. To over-come the problem of isolating $a_{PHY}(\lambda)$ from $a_{SES}(\lambda)$ Kishino *et al.* (1985) firstly measured $a_{SES}(\lambda)$ on a filter pad and then extracted the phytoplankton pigments in methanol, thereby measuring $a_{TRI}(\lambda)$. *In vivo* $a_{PHY}(\lambda)$ was calculated as the difference between the two measurements. The most common analytical techniques employed to over-come the problem of particulate scattering involve the use of a spectrophotometer fitted with an integrating sphere which serves to collect a larger proportion of the scattered light than

do the traditional entrance slits to the spectrophotometer's photomultipliers (Kirk 1981a, Dekker 1993). Alternatively the 'opal glass' method can be used whereby the collimated sample and reference beams of the spectrophotometer are diffused by rotating both cuvettes through 90°. Diffusing the beams has the affect of increasing the reception angle of the photomultipliers thereby allowing a greater proportion of the scattered, but transmitted, light to be recorded (Shibita 1958, Davis-Colley *et al.* 1986, Sathyendranath *et al.* 1987). Finally the 'filter pad' method can be used whereby a spectrophotometer is used to measure the absorption of seston collected upon a glass fibre filter pad held close to the entrance port to the spectrophotometer's photomultiplier (Kishino *et al.* 1985, Bricaud and Stramski 1990). A comparative review of these various techniques was conducted by Maske and Haardt (1987). The concentration problem was addressed by Malthus and Dekker (1990) who concentrated the particulate fraction of a large water sample on a membrane filter from which it was possible to resuspend, through gentle agitation, the particulates in a reduced volume of water.

To compare variations in $a_{PHY}(\lambda)$ from phytoplankton at different concentrations Beer's Law is used to calculate the specific absorption $a_{PHY}^*(\lambda)$ (m^2 [mg CHL*a*]$^{-1}$) (Morel and Bricaud 1981, Hoepffner and Sathyendranath 1991):

$$a_{PHY}^*(\lambda) = a_{PHY}(\lambda) / \text{mg } CHLa \text{ } L^{-1} \quad \text{.....2.1.5.2.2}$$

CHL*a* is usually chosen for this normalisation because it is both ubiquitous and is frequently measured as an index of phytoplankton biomass (Hoepffner and Sathyendranath 1991). Assuming that phytoplankton cells have a consistent ratio between cell volume and CHL*a* concentration, calculating $a_{PHY}^*(\lambda)$ should remove all variation in $a_{PHY}(\lambda)$ that can be attributed to biomass concentration, therefore $a_{PHY}^*(\lambda)$ provides a more suitable basis for comparing spectra in terms of pigment composition rather than pigment concentration. $a_{PHY}^*(\lambda)$ is also needed for a meaningful interpretation of the actual optical properties of waters with respect to their phytoplankton content and to relate bio-optical properties of a water body to its phytoplankton content (Morel and Bricaud 1981, 1986).

After variations in pigment composition the main influence on $a_{PHY}(\lambda)$ is the package effect (also known as the particle effect or the sieve effect) (Kirk 1975). The absorption spectrum of any suspension of strongly absorbing particles will appear flattened compared to the absorption of a solution containing the same molecules in homogenous dispersions: the stronger the absorption the stronger the flattening influence of the package effect (Duysens 1956). The package effect arises because of the non-linear relationship between absorption efficiency and pigment concentration or cell size, so that increases in pigment concentration or increases in cell size induce diminishing increases in absorption efficiency (Nelson and Robertson 1993). In the case of $a_{PHY}(\lambda)$ the diminishing increases in absorption efficiency arises from the enhanced mutual shading of both the pigmented molecules held within the discrete thylakoids present in the cell chloroplasts and the shading of one cell by another (Duysens 1956, Hoepffner and Sathyendranath 1991). The degree to which the package effect

influences $a_{PHY}(\lambda)$ is directly related to the absorption characteristics of the cell and the cell morphology. The most inefficient cell shape, in terms of absorption, is spherical, while the most efficient cell form is long, thin and cylindrical. Furthermore the smaller the cell (the more it approaches the morphology of an equivalent solution) the greater the absorption efficiency of each cell (Kirk 1994).

To compare *in vivo* $a_{PHY}(\lambda)$ from different classes on the basis of pigment content alone the particle effect must first be removed (Sathyendranath *et al.* 1987). Theoretically the package effect can be removed by considering the absorption by an equivalent solution $a_{SOL}(\lambda)$. Differences highlighted by the comparison of $a_{SOL}(\lambda)$ for different phytoplankton classes should be entirely due to pigment differences between those classes (Morel and Bricaud 1983). Sathyendranath *et al.* (1987) provide the equations for the calculation of $a_{SOL}(\lambda)$.

Phytoplankton scattering $b_{PHY}(\lambda)$ is magnitudinally more important than $a_{PHY}(\lambda)$ (Davis-Colley *et al.* 1986). $b_{PHY}(\lambda)$ tends to exhibit simple spectral forms being either inversely proportional to wavelength or exhibiting no variation with wavelength with the exception of those $b_{PHY}(\lambda)$ features induced by anomalous dispersion in the proximity of pigment absorption bands (Davis-Colley *et al.* 1986, Morel and Bricaud 1986). Cells with gas vacuoles scatter more light than those without (Ganf *et al.* 1989). This can be important as the occurrence of gas vacuoles is restricted to class Cyanophyceae. Phytoplankton cells are considered to be optically large and have low refractive indices hence they predominantly scatter at large forward angles with some backscattering (Doucha and Kubin 1976). Ganf *et al.* (1989) showed that the presence of gas vacuoles in *M. aeruginosa* caused a ten times increase in scatter at 90° and a five times decrease in forward scatter.

2.1.6 Optical efficiency factors

The effectiveness of absorption or scattering exhibited by an optically active water quality parameter is described by its optical efficiency factors. The optical efficiency factors are absorption efficiency Q_a, scattering efficiency Q_b and attenuation efficiency Q_c (Morel and Bricaud 1986). Optical efficiency factors are dimensionless and are independent of particle concentration. They correspond to the ratios of energy impinging upon the particle to the energy 'affected' by that particle and are considered with respect to the hypothetical black body.

Phytoplankton cells have evolved to be extremely efficient at absorbing light in an aquatic medium (C.S.Reynolds pers. comm.). This efficiency comes from the design of their external morphology, their internal cell structure and their pigment composition (Britton 1983).

$$a_{TOT}(\lambda) = a_{H2O}(\lambda) + a_{DYS}(\lambda) + a_{PHY}(\lambda) + a_{TRI}(\lambda) \quad \text{2.1.7.1.}$$

Similarly for the total scattering coefficient $b_{TOT}(\lambda)$:

$$b_{TOT}(\lambda) = b_{H2O}(\lambda) + b_{DYS}(\lambda) + b_{PHY}(\lambda) + b_{TRI}(\lambda) \quad \text{2.1.7.2}$$

The sum of $a_{TOT}(\lambda)$ and $b_{TOT}(\lambda)$ is the total attenuation coefficient $c_{TOT}(\lambda)$:

$$c_{TOT}(\lambda) = a_{TOT}(\lambda) + b_{TOT}(\lambda) \quad \text{2.1.7.3}$$

The ratio of $a_{TOT}(\lambda) : b_{TOT}(\lambda)$ describes the probability of whether a photon is likely to be scattered or absorbed. Kirk and Tyler (1986) note that attenuation of light within the water column is primarily due to absorption. However, scattering can greatly increase the number of absorption events taking place by increasing the pathlength of individual photons.

2.2 The apparent optical properties of natural water bodies

It has been established that the inherent optical properties of a water body are a function of the composition of the optically active water quality components, however their inherent optical properties are not easily measured. Apparent optical properties are those optical properties of a water body that are a function of both the optically active components of the water body and the illuminating light field. It is easier to derive information pertaining to the concentration of each water quality parameter from the apparent optical properties because typically they are easier to measure. The apparent optical properties that are most commonly measured are water colour, downwelling attenuation $Kd(\lambda)$ and reflectance $R(\lambda)$. Measurements of apparent optical properties are more easily made in the field than the inherent optical properties but, because they are a function of the incident light field and because they are a combination of the absorption and scattering properties of each component, they can be more difficult to interpret.

Water colour is usually measured by reference to the standard colours of the Hazen system (Standard Methods, 1992). $Kd(\lambda)$ can be measured by using a submersible light sensor (Malthus and Dekker 1990), but it is more common to use a Secchi disk as a surrogate measure of attenuation over the visible spectrum. Secchi disk depth describes water clarity which is a function of the attenuation across the visible spectrum, Kd(PAR) where PAR is Photosynthetically Active Radiation (Effler *et al.* 1988). Kd(PAR) is used to define the euphotic depth which is generally given as the depth at which downwelling irradiance is reduced to 1% of its surface value. Euphotic depth defines the depth to which aquatic plants can generally flourish (Kirk 1994). For the purpose of assessing water quality reflectance is measured radiometrically. Measurement of this particular apparent optical property is discussed in more detail below.

2.2.1 Generation of the water-leaving upwelling radiance flux

A photon incident at the surface of a water body will either be reflected at the air-water interface or it will penetrate into the water. Surface reflected radiance $Lr_a(\lambda)$ contains no information pertaining to the water quality but can interfere with the measurement reflectance. $Lr_a(\lambda)$ is a function of sun angle

and wave action (Dekker 1993) and is dominated by photons from the diffuse skylight due to the more obtuse angles of diffuse source photons. As skylight is predominantly blue due to atmospheric Raleigh scattering (Monteith and Unsworth 1990) $Lr_a(\lambda)$ is biased to the shorter wavelengths. Surface reflection is governed by Fresnel's equation for reflectance (Baker and Smith 1990).

Photons that are transmitted through the air-water interface undergo refraction which is described by Snell's Law (Kirk 1994). Considering optically deep waters (where the influence of bottom reflectance can be ignored) there are two potential fates for these downwelling photons, they will either be absorbed by one of the optically active water quality parameters or they will be scattered back through the water-air interface. The probability of either of these two events occurring is determined by the relative magnitudes of the inherent optical properties $a_{TOT}(\lambda)$ and $b_{TOT}(\lambda)$. Scattering of downwelling photons changes their direction in accordance with $\beta(\theta)$. Multiple forward scattering events or single backscattering events will generate an upwelling flux $Lu_w(\lambda)$. If $b_{TOT}(\lambda)$ is high then the percentage of $Ld_w(\lambda)$ converted into $Lu_w(\lambda)$ will also be high. In addition, because some of the optically active components have scattering coefficients which are wavelength dependent, the generated $Lu_w(\lambda)$ can become biased towards specific wavelengths. Absorbed photons are removed from the underwater light climate hence they can not become part of the $Lu_w(\lambda)$ flux and, because the optically active components discussed above preferentially absorb at specific wavelengths, $a_{TOT}(\lambda)$ further biases the spectral quality of $Lu_w(\lambda)$. The combined inherent optical properties therefore determine the spectral quality of the $Lu_w(\lambda)$.

Some of the photons which become part of the $Lu_w(\lambda)$ flux will reach the water-air interface where again they may be reflected, this time back down into the water body, the remainder will be refracted as they are transmitted through the interface. The $Lu_w(\lambda)$ flux which passes out from the water's surface becomes the water-leaving upwelling flux $Lu_a(\lambda)$. Because $Lu_a(\lambda)$ has been spectrally altered according to the inherent optical properties of the optically active water quality components it contains information pertaining to those components. It is this flux that imparts colour to the water body as perceived by the eye and it is this flux that is measured by remote sensors monitoring water quality (Kirk 1994).

2.2.2 Fluorescence and bottom reflectance

Some of the optically active water quality components lose the electromagnetic energy gained during absorption by fluorescence (re-emitting light at a longer wavelength). The phytoplankton pigment CHL*a* fluoresces at 685 nm and PCE fluoresces at 580 nm (Britton 1989, Lee *et al.* 1994). Various forms of DOC also fluoresce in the visible spectrum (Lee *et al.* 1994). Fluoresced photons can be emitted in any direction so there is a 50% chance that fluoresced photons will join the upwelling flux $Lu_w(\lambda)$ where they will be detected by remote sensing instruments (Topliss and Platt 1986).

In optically shallow water bodies photons may reach the lake bed where, depending upon the nature of that bed, they will either be absorbed or scattered. Absorbed photons are lost from the underwater light climate while scattered photons will join $Lu_w(\lambda)$. Consequently bottom reflectance, which is a function of the absorption properties of the lake bed, has implications for the spectral quality of $Lu_w(\lambda)$ as measured by remote sensing instruments.

2.2.3 The measurement and calculation of reflectance

$Lu_a(\lambda)$ is not only a function of the inherent optical properties of a water body but is also a function of the continuously varying surface incident solar irradiance (Jackson *et al.* 1987). To overcome this dependency on incident irradiance the spectral signature, or reflectance $R(\lambda)$, is calculated. $R(\lambda)$ allows multitemporal comparisons of the optical properties of the same water body or comparisons of different water bodies to be made. Various forms of spectral signature can be calculated according to the location and direction of sensors and according to sensor type (Dekker 1993).

For the purpose of calculating reflectance a measurement of the water-leaving upwelling radiance $Lu_a(\lambda)$ must be performed in combination with a reference measurement of the downwelling surface incident flux. This reference measurement $Lo(\lambda)$ is usually performed using a narrow angle field-of-view (FOV) viewing in the nadir over a reference panel with Lambertian properties held perpendicular to the nadir (Milton 1987). In reality $Lo(\lambda)$ must be corrected for the non-Lambertian properties of the panel (Jackson *et al.* 1987). The target radiance measurements made when viewing in the nadir with a restricted angle FOV will be of the above surface upwelling flux $Lz_a(\lambda)$ which is a combination of $Lu_a(\lambda)$ and $Lr_a(\lambda)$.

Apparent above surface radiance reflectance is measured +0.01 m above the surface of the water $R(\lambda)(+0.01\ m)$ (Dekker 1993). $R(\lambda)(+0.01\ m)$ is also known as bi-directional reflectance (Milton 1987, Jackson *et al.* 1987) and is calculated according to:

$$R(\lambda)(+0.01\ m) = Lz_a(\lambda) / Lo(\lambda) \quad \text{.....2.2.3.1}$$

However $Rz_a(\lambda)$ is based upon $Lz_a(\lambda)$ which includes the upwelling components of $Lu_a(\lambda)$ and $Lr_a(\lambda)$. Because $Lr_a(\lambda)$ is dependent both on the time of day (the angle of incidence of the direct solar beam determines the degree of surface reflectance) and the sky conditions (cloud cover and aerosol concentration modify the diffuse flux) and wave action, and because it contains no information pertaining to the water quality, the component of $Lr_a(\lambda)$ present in measured $Lz_a(\lambda)$ reduces the comparability of reflectance measurements calculated according to Equation 2.2.3.1 which have not taken $Lr_a(\lambda)$ into consideration (Dekker 1993).

A more appropriate measure of reflectance for multitemporal comparison purposes is sub-surface irradiance reflectance $R(\lambda)(-0.01m)$ which is reflectance measured at a depth of -0.01 m, or just below the water's surface (Dekker 1993). To obtain $R(\lambda)(-0.01m)$ additional field measurements are required

to characterise the diffuse and direct components of the incident flux ($Ed_{DIF}(\lambda)$ and $Ed_{SUN}(\lambda)$). These measurements are of the diffuse flux $Ld_{aDIF}(\lambda)$ and are performed with the a narrow FOV viewing in the zenith and with a similar FOV viewing at 90° to the zenith angle of the direct incident solar flux and 180° to the azimuth angle of that flux. This latter location identifies the $Ed_{DIF}(\lambda)$ minima (Dekker 1993). $F(\lambda)$ is the fraction of $Ed_a(\lambda)$ that is diffuse and can be estimated from:

$$F(\lambda) = Ld_{aDIF}(\lambda) / Lo(\lambda) \quad \text{2.2.3.2}$$

Knowing $F(\lambda)$, $Lo(\lambda)$ and solar zenith angle at the time of measurement (which is used along with the refractive indices of air and water to calculate Fresnel reflectance) allows calculation of the fraction of the downwelling flux that is reflected from the surface of the water that will form $Lr_a(\lambda)$. Assuming that diffuse skylight has a uniform radiance distribution, then 6.6% of surface incident radiation will be reflected from the water's surface, however wind induced surface waves will alter this percentage (Baker and Smith 1990). $Lr_a(\lambda)$ can then be removed from $Lz_a(\lambda)$ leaving $Lu_a(\lambda)$. $Lu_a(\lambda)$ can then be converted to sub-surface upwelling radiance $Lu_w(\lambda)$ by taking into account the effect of refraction at the air-water interface according to:

$$Lu_w(\lambda) = 1.84\ Lu_a(\lambda) \quad \text{2.2.3.3}$$

following Dekker (1993). This can be converted into upwelling sub-surface irradiance $Eu_w(\lambda)$ by multiplication by the Q factor (the ratio of irradiance to radiance underwater, $Q = Eu_w / Lu_w$). Were water to have Lambertian properties Q would equal π, however the underwater light climate is not Lambertian (Kirk 1994) and Q has been variously measured to lie between 1 and 5 (Dekker 1993). The influence of Q on the calculation of reflectance is discussed in Chapter Four, but as long as the magnitude of Q remains the same at all wavelengths, then its value will have no spectral influence on reflectance. Above-surface irradiance reflectance $R(\lambda)(+0.01m)$ (reflectance at +0.01 m above the water's surface) can be calculated according to:

$$R(\lambda)(+0.01m) = Eu_w(\lambda) / Ed_a(\lambda) \quad \text{2.2.3.4}$$

The $Lo(\lambda)$ measurement can then be converted into above-surface downwelling irradiance $Ed_a(\lambda)$ by multiplying by π and taking into account the non-Lambertian properties of the reference panel. This is then converted to sub-surface irradiance reflectance $Ed_w(\lambda)$ by taking into account the fraction of $Ed_a(\lambda)$ that is lost to Fresnel reflectance and including the contribution to $Ed_w(\lambda)$ from that part of $Eu_w(\lambda)$ that is affected by Fresnel reflectance. For completely overcast skies $Ed_a = 0.94Ed_w$ with this fraction increasing as the incident flux becomes more direct (Baker and Smith 1990). Sub-surface irradiance reflectance $R(\lambda)(-0.01m)$ can then be calculated according to:

$$R(\lambda)(-0.01m) = Eu_w(\lambda) / Ed_w(\lambda) \quad \text{2.2.3.5}$$

The benefit of $R(\lambda)(-0.01m)$ is that as many of the factors which reduce the appropriateness of comparison of $R(\lambda)(+0.01$ m) have been taken into consideration (Dekker 1993). There are assumptions made in the calculation of both $Eu_w(\lambda)$ from $Lz_a(\lambda)$ and $Ed_w(\lambda)$ from $Lo(\lambda)$. These assumptions may not always be valid, so the suitability of the $R(\lambda)(-0.01m)$ over $R(\lambda)(+0.01$ m) needs to be taken into account.

Due to its dependence upon the underwater upwelling radiance flux, reflectance is directly related to the inherent optical properties of a water body. This was theoretically shown by Gordon *et al.* (1975) who stated that:

$$R(\lambda) = r_o + r_1(b_{bTOT}(\lambda) / (b_{bTOT}(\lambda) + a_{TOT}(\lambda))) + r_2(b_{bTOT}(\lambda) / (b_{bTOT}(\lambda) + a_{TOT}(\lambda)))^2 \quad ...2.2.3.6$$

where r_n are constants. This polynomial shows that $R(\lambda)$ varies directly with $b_{bTOT}(\lambda)$ and inversely with $a_{TOT}(\lambda)$. A similar relationship was proposed by Morel and Prieur (1977) who found that Gordon's polynomial relationship could be reduced to a linear expression if the term $b_{bTOT}(\lambda) / (b_{bTOT}(\lambda) + a_{TOT}(\lambda))$ was replaced by $b_{bTOT}(\lambda) / a_{TOT}(\lambda)$, giving:

$$R(\lambda) = r_1 (b_{bTOT}(\lambda) / a_{TOT}(\lambda)) \quad .....2.2.3.7$$

where $R(\lambda)$ was surface irradiance reflectance ($R(\lambda)(+0.01$ m) and r_1 was found to be approximately equal to 0.33. Note that r_1 changes to 0.1076 when $R(\lambda)$ equals sub-surface irradiance reflectance ($R(\lambda)(-0.01$ m)) because surface effects and the Q factor require consideration (Carder *et al.* 1986). Bearing in mind that the inherent optical properties are strictly additive, the influence of each of the various optically active water quality components on reflectance could be calculated.

If the spectral nature of $b_{bTOT}(\lambda)$ is non-trending, or has a simple linear form, then according to Equation 2.2.3.7 all the spectral variations in $R(\lambda)$ are the result of the variations in $a_{TOT}(\lambda)$. This being the case, the various components of water quality that contribute to $a_{TOT}(\lambda)$ as calculated in Equation 2.1.7.1. should all feature in $R(\lambda)$. It is this relationship that allows information pertaining to the quality and quantity of these optically active components to be extracted from $R(\lambda)$ (Morel and Prieur 1977, Kitchen and Zaneveld 1990). Measurements of water quality from $R(\lambda)$ should be robust even in the presence of spectral variations in $b_{bTOT}(\lambda)$ which, according to Equation 2.2.3.7, is also responsible for determining $R(\lambda)$. This is because spectral changes in $a_{TOT}(\lambda)$ tend to be more significant than changes in $b_{TOT}(\lambda)$ for the different optically active water quality components discussed above (Kitchen and Zaneveld 1990).

It is because $R(\lambda)$ is determined by a combination of the inherent optical properties of the various optically active water quality components, that the interpretation of spectral variations in $R(\lambda)$ is difficult. Furthermore, because the variations in $R(\lambda)$ attributable to individual water quality components can be difficult to isolate from each other, when using remote sensing to measure a specific water quality component the spectral variation attributable to that component should not be considered in isolation from the spectral effects attributable to the other components. To overcome this problem a technique has been employed whereby, in a sample obtained from the water body, the various water quality components are physically isolated from each other. The inherent optical properties of each parameter can then be measured separately thereby allowing a more comprehensive understanding of the spectral signature (Dekker *et al.* 1992a). Breaking-down $R(\lambda)$ into its various

components is a useful technique when trying to develop new remote sensing algorithms for the retrieval of information pertaining to water quality, or in this instance to phytoplankton type.

2.3 Using airborne remote sensors to monitor inland water quality

2.3.1 The advantages of airborne over satellite sensors for monitoring freshwater lakes

The above section has shown how the information pertaining to the water quality is contained in the various forms of a water body's spectral signature as measured by remote sensing. Due to the low levels of solar UV light at surface levels and the strong IR absorption properties of waters, the useful information in the spectral signature relating to water quality is restricted to the visible wavelengths (400 to 700 nm). In the presence of strong absorption by coloured DOC the region of the spectral signature containing information relating to phytoplankton type and concentration is further restricted to the green and the red wavelengths (500 to 700 nm).

Satellite borne sensors have been used for the monitoring of lake water quality (Almanza and Melack 1985) but for remote sensing inland water bodies airborne instruments have certain advantages over satellite platforms when considered in terms of spectral, spatial, radiometric and temporal resolution (Malthus *et al.* 1996). The FLI/PMI was developed specifically for the mapping of CHL*a*, it measures from 430 to 805 nm recording at 8 wavebands only (Curran 1994). MERIS and MODIS, two satellite sensors, record in 15 and 19 wavebands over the electromagnetic spectrum from 400 to 1050 nm (Curran 1994). The newest developments in airborne sensors (*e.g.* the Compact Airborne Spectrographic Imager or CASI) provide hyperspectral resolution throughout the visible wavelengths (Zacharias 1991). When operated in spectral mode the CASI samples up to 228 wavebands over the range 410 to 925 nm. This compares favourably with the FLI/PMI, ATM and the CAESAR sensors which are the other airborne sensors typically used for aquatic studies (Zacharias 1992, Dekker 1993).

The spatial resolution of satellite sensors is traditionally low (Millie *et al.* 1992), with the exception of SPOT HRV which achieves 20 - 30 m (Malthus *et al.* 1996). The spatial resolution of low altitude airborne sensors is much more suitable to the remote sensing of the relatively small (in terms of their spatial extent) freshwater lakes of the UK. Satellite sensors are programmed to cope with a very broad range of radiometric magnitudes however, water bodies are relatively poor reflectors compared to land surfaces, consequently satellite images of lakes generally lack contrast (Malthus *et al.* 1996). This means that the ratio based algorithms for the retrieval of water quality parameters rely on a limited range of Digital Number (DN) values giving them a low signal to noise ratio (Dekker and Peters 1993).

Satellites are restricted in their orbits so their return periods to a particular water body are restricted (*e.g.* LANDSAT's return period is 16 days (Millie *et al.* 1992)). Moreover there is no guarantee that

the target water body will not be obscured by cloud. Airborne sensors can be deployed with much greater flexibility taking advantage of clear skies. This temporal flexibility is important for the remote sensing of phytoplankton blooms which can develop and decline within a matter of weeks (George 1993). Furthermore airborne sensors are particularly useful for the study of lake processes because images can be obtained on consecutive days (Dekker *et al.* 1992a). Consequently, because of their improved spatial, spectral and radiometric resolution and because of their greater deployment flexibility, the airborne platform is the preferred platform for remote sensing inland waters

2.3.2 Examples of remote sensing campaigns for monitoring freshwater lakes

George and Hewitt (1990) used the DAEDALUS Airborne Thematic Mapper to measure phytoplankton concentrations in some of the lakes of the English Lake District. Dekker *et al.* (1991) used the Programmable Imager of Moniteq Inc. (PMI) to study the temporal and spatial variations in water quality (measured in terms of biological, chemical and physical parameters) of the Loosdrecht Lakes, the Netherlands. These lakes represent a good cross-section of optical characteristics including both Case I and II water types. Harding *et al.* (1992) described the use of NASA's airborne remote sensor Ocean Data Acquisition System (ODAS) for the purposes of measuring the distribution of phytoplankton biomass in the Chesapeake Bay. They found that the airborne sensor offered improved temporal and spatial resolution over satellite systems which lead to a better understanding of the phytoplankton dynamics in the estuary. Millie *et al.* (1992) used the Calibrated Multispectral Scanner (CAMS) to monitor the concentration of the two classes of phytoplankton from their pigment concentrations in the Case II waters of fish aquaculture impoundments in Mississippi, USA. Ramsey (1992) used the Airborne Thematic Mapper (ATM) multi-spectral scanner to study the water quality in reservoirs used for cooling water. Gitelson *et al.* (1993) also note that inland waters, with their Case II optical properties, present particular problems to the remote sensing of water quality. Dekker *et al.* (1991) present a series of algorithms for the purpose of monitoring various water quality parameters. These algorithms were principally developed from hyperspectral data from the Case II waters of the Loosdrecht Lakes, Holland, as measured by the CASI sensor. Lee *et al.* (1994) used the Airborne Visible and Infrared Imaging spectrometer (AVIRIS) to assess the performance of their modelling exercise for Case II waters.

Although remote sensing has been used to monitor a variety of water quality parameters the most frequently monitored water quality parameter is CHL*a*, the others being Secchi disk depth, suspended solids concentration and coloured DOC (Carder *et al.* 1989, Zacharias *et al.* 1992, Dekker *et al.* 1992a). Typically reflectance algorithms for the monitoring of water quality tend to be ratio-based algorithms which have been established by linear regression (Andre 1992). Such algorithms work well in Case I waters where reflectance is primarily a function of the phytoplankton concentration because the other optically active water quality components are present at relatively low concentrations. Ratio based algorithms are less accurate in Case II waters where optical characteristics are determined be a mixture of optically active water quality components which combine to produce a reflectance spectra

that can be independent of phytoplankton concentration (Lee *et al.* 1994). In particular, the influence of tripton on reflectance means that algorithms developed for oceanic Case I waters (where tripton is essentially absent) are often inappropriate for Case II inland waters where the tripton component can be significant (Gitelson 1993). Furthermore, ratio-based CHL*a* retrieval algorithms developed by workers monitoring phytoplankton in inland Case II water bodies with high concentrations of coloured DOC, tend to use wavelengths located in the red wavelengths where absorption by phytoplankton is less significant than in the Soret band, but the conflicting absorption signal from DYS is minimal (Bukata *et al.* 1985, Dekker 1993).

Remote sensors of water quality generally treat water bodies as homogeneous in their vertical extent and do not attempt to account for variations in water quality arising from vertical stratification. Modellers of the under-water light climate have considered the vertical depth of water bodies. Andre (1992) modelled the effects of a vertically stratified phytoplankton biomass on reflectance and found them to be magnitudinal rather than spectral. Gordon and MacCluney (1975) used Monte-Carlo simulations of radiative transfer theory in homogenous oceanic waters to determine that the depth attained by 90% of the reflected water-leaving photons can be described by the reciprocal of the downwelling attenuation coefficient Kd(λ). Therefore measurement of R(λ) should provide information pertaining to the quality of the water column to this depth.

2.3.3 Field radiometric measurements

Field spectroscopy has three important roles to play in remote sensing (Milton 1987). Firstly, field radiometric measurements of upwelling radiance are performed in the absence of atmospheric interference which can assist in the calibration of airborne data. Secondly, field measurements are useful when configuring airborne sensors in terms of selecting bands, optimal flight times and viewing configuration. Thirdly, because field radiometers such as the Macam SR9910 and the Spectron SE 590 (introduced in § 4.3.1.1 and § 5.2.1 respectively) have hyperspectral capabilities they provide input data for the development of algorithms for the prediction of the concentrations of various water quality parameters (Dekker *et al.* 1992a, Gitelson 1993, George and Charlton 1996).

As mentioned above, field spectroradiometers and some airborne imagers, perform high spectral resolution radiometric measurements. High spectral resolution undoubtedly offers potential for improving indices currently used for the remote sensing of aquatic environments. However Price (1991) noted that increasing the spectral resolution of reflectance spectra from a few broad bands to many narrow bands (hyperspectral data) will not provide a linear increase in data yield. Many spectrally close data points will be highly correlated with each other and in their response to changes in water quality parameters. Similarly Gitelson *et al.* (1993) used factor analysis to show that the hyperspectral reflectance spectra from various inland water bodies at a range of trophic status could be effectively summarised by three eigenvectors explaining 90% of the variation in measured reflectance. The first eigenvector highlighted the importance of absorption by phytoplankton and coloured DOC in

the blue wavelengths, the second highlighted the influence of phytoplankton on the green peak and the third highlighted the near-IR phytoplankton reflectance peak. Consequently, increasing the number of bands will not lead to a linear increase in the information retrieval, and reducing the large data set to a manageable data set without loss of information is one of the main problems associated with hyperspectral data sets. Price (1991) noted that hyperspectral data sets do not always provide significant increases in information over that provided by broad band data sets because spectrally close data points will be highly correlated with each other. This not-with-standing the availability of hyperspectral reflectance spectra does facilitate the optimisation of the spectral location of the broader bands, the development of the most appropriate algorithms for retrieval of water quality information, the development of new techniques for the retrieval of that information and leads to a more comprehensive understanding of the spectral signature.

Although hyperspectral field measurements of aquatic targets are undoubtedly useful they are hampered by factors which are essentially beyond the control of the field-worker. These factors introduce uncertainty to the field measurements consequently hindering subsequent analysis and interpretation of the data. In the field, the prevailing irradiance conditions under which radiometric field measurements are conducted are subject to the flux and spectral inconsistencies due to variable atmospheric conditions which pose a persistent problem for the measurement of $R(\lambda)$ in temperate climates. Added to this is the dynamic nature of the aquatic targets which vary both temporally and spatially due to compositional changes in the water quality (Prezlin and Boczar 1986). The drawbacks of the field environment are:

- the angle of incident solar radiance varies diurnally and annually as the earth moves on its axis and orbits the sun. This variation has implications for the calculation of the spectral signature (Dekker 1993);
- variations in the cloud cover cause temporal changes in the incident solar flux. Unless they are measured simultaneously, changes in radiance flux that occur between the reference measurement and the target measurement lead to errors in the calculation of the spectral signature;
- temporal changes in the incident radiation requires the employment of fast CCD array type spectroradiometers (*e.g.* the Spectron SE-590) to obtain a spectral signature. These radiance changes preclude the use of the more accurate and precise, but slower, motor driven monochromatic photo-multiplier type spectroradiometers (*e.g.* the Macam SR 9910);
- the ratio of diffuse to direct irradiance (Equation 2.2.3.2) has implications for surface reflectance and is a function of atmospheric conditions. The water surface will specularly reflect incident photons with a trajectory zenith angle $< 45°$ as described by the Fresnel equation;
- water waves change the angle of interaction between the upwelling radiance and the water surface. This has implications for the amount of upwelling radiance flux reflected back into the water body as described by the Fresnel equation and for the angle of the water-leaving upwelling reflectance;

- the viewing angle of a boat-mounted sensor changes as the boat rocks with water waves. This has implications for the consideration of $Lr_a(\lambda)$;
- the constituent composition of aquatic targets can alter during long field sessions, this reduces the appropriateness of corroborative spot samples of water quality used to ground truth the field radiometry;
- the optically active water quality components of the aquatic target are complex. Even Case I waters have spectral signatures which contain spectral information from the water molecules, from fine suspended sediment particles, from dissolved organic matter, and from the range of phytoplankton. The Case II water bodies have even more complex spectral signatures. The ability to identify phytoplankton class from reflectance spectra could therefore be compromised because of the difficulty in isolating the phytoplanktonic reflectance signal from the signals from the other optically active water quality parameters.

Thus, if spectral signatures are to be measured to the level of accuracy and precision required for the development of phytoplankton class identification routines, the adverse factors mentioned above must be controlled. This can be achieved by setting up a scaled aquatic environment in the laboratory (Mantovani and Cabral 1992, Novo *et al.* 1991, Quibell 1991, 1992). During this study an experimental laboratory based tank was constructed for this purpose which is described in Chapter Four.

2.4 Chapter Two summary

The optical quality of natural waters is controlled by the inherent optical properties of the optically active water quality components. In Case I waters this essentially means the apparent optical property of reflectance is a function of the inherent optical properties of the scattering and absorption by water and phytoplankton. In Case II waters the reflectance signal can be more complex being a function of a variety of optically active water quality components.

It has been shown that, in the absence of significant variations in $b_{bPHY}(\lambda)$, $R_{PHY}(\lambda)$ is the inverse of $a_{PHY}(\lambda)$. This leads to the hypothesis that measurements of $a_{PHY}(\lambda)$ can be used as surrogates for measurements of $R_{PHY}(\lambda)$, hence a database of pure-culture phytoplankton absorption spectra could be used for the development of algorithms and routines for the identification of phytoplankton class, which could then be applied to pure-culture phytoplankton reflectance spectra. It has also been shown that phytoplankton class can be distinguished on the basis of photosynthetic pigment composition and because pigment absorption modifies the spectral quality of the upwelling water-leaving reflectance there is potential to determine pigment composition, and therefore the identity of phytoplankton class through the remote sensing.

The theory and techniques behind field and aerial remote sensing have been introduced. In particular the problems caused by surface effects at the air-water interface have been discussed and the need to obtain a measurement of the most relevant form of reflectance, sub-surface irradiance reflectance, was highlighted. The techniques for calculating the $R(\lambda)(-0.01\ m)$ from field spectra have been introduced.

Previous remote sensing campaigns aimed at monitoring the quality of inland water bodies have been described. These studies have mainly dealt with measuring the concentration of phytoplankton blooms. Few have tackled the problem of phytoplankton class identification. The benefits of hyperspectral data with its increased potential for information retrieval over traditional broad band sensors has been shown.

This review suggests that, to investigate the potential for identifying phytoplankton class from remotely sensed data, an experimental approach would be appropriate. This would allow control of a range of extraneous factors which can adversely affect field measurements.

Chapter Three Experimental Design

No chapter deals exclusively with experimental methodology because each progressive chapter has distinct methods which are presented accordingly. This chapter explains the fundamental approach to experimental design used throughout this study.

Chapter One has introduced the concept of phytoplankton identification through measurement of reflectance spectra and has detailed why the ability to monitor the type of phytoplankton dominating a crop through remote sensing might be desirable. Chapter Two has presented the theory behind the optical properties of the various optically active water quality parameters present in water bodies and has shown how reflectance is a function of the various inherent optical properties.

For this study the key optical relationships are those between phytoplankton pigment composition and phytoplankton absorption spectra and those between absorption and reflectance spectra. To summarise: phytoplankton's photosynthetic pigment compositions are essentially similar for species from the same class; $a_{PHY}(\lambda)$ is primarily a function of pigment composition; and $a_{PHY}(\lambda)$ is essentially the inverse of $R_{PHY}(\lambda)$. The experimental design of this study was developed from these two relationships. Although $b_{PHY}(\lambda)$ can influence both the magnitude and spectral shape of $R_{PHY}(\lambda)$ it was established in Chapter Two that $R_{PHY}(\lambda)$ is primarily a function of $a_{PHY}(\lambda)$. Consequently a database of measured pure culture high-spectral $a_{PHY}(\lambda)$ was assembled to develop algorithms for the qualification of phytoplankton type. These algorithms were then to be applied to pure culture high-spectral $R_{PHY}(\lambda)$ as measured in a specially designed experimental tank in the controlled conditions of the laboratory. The algorithms were then be tested on field-based reflectance spectra made at the surface of a naturally eutrophic water body. Finally the identification algorithms were tested on reflectance spectra for the same water body as measured at altitude. The advantage of this experimental design is that pure culture $a_{PHY}(\lambda)$ are more easily obtained in the laboratory than pure culture $R_{PHY}(\lambda)$ consequently the algorithms could be developed from a more extensive set of measurements (in terms of freshwater phytoplankton types and classes). This approach should mean that the developed phytoplankton identification algorithms are more robust, thus able to cope with the identification of freshwater phytoplankton from the spectral signatures of a wider range of lakes and from the same lake throughout the growing season. Understanding the relationships between $a_{PHY}(\lambda)$, $b_{PHY}(\lambda)$ and $R_{PHY}(\lambda)$, combined with the availability of pure culture $R_{PHY}(\lambda)$ also assisted in the development and testing of quantitative algorithms for the monitoring of concentration of the phytoplankton crop. In particular the affects of phytoplankton type on the precision of these quantitative algorithms was be tested. The experimental design is summarised in Figure 2.

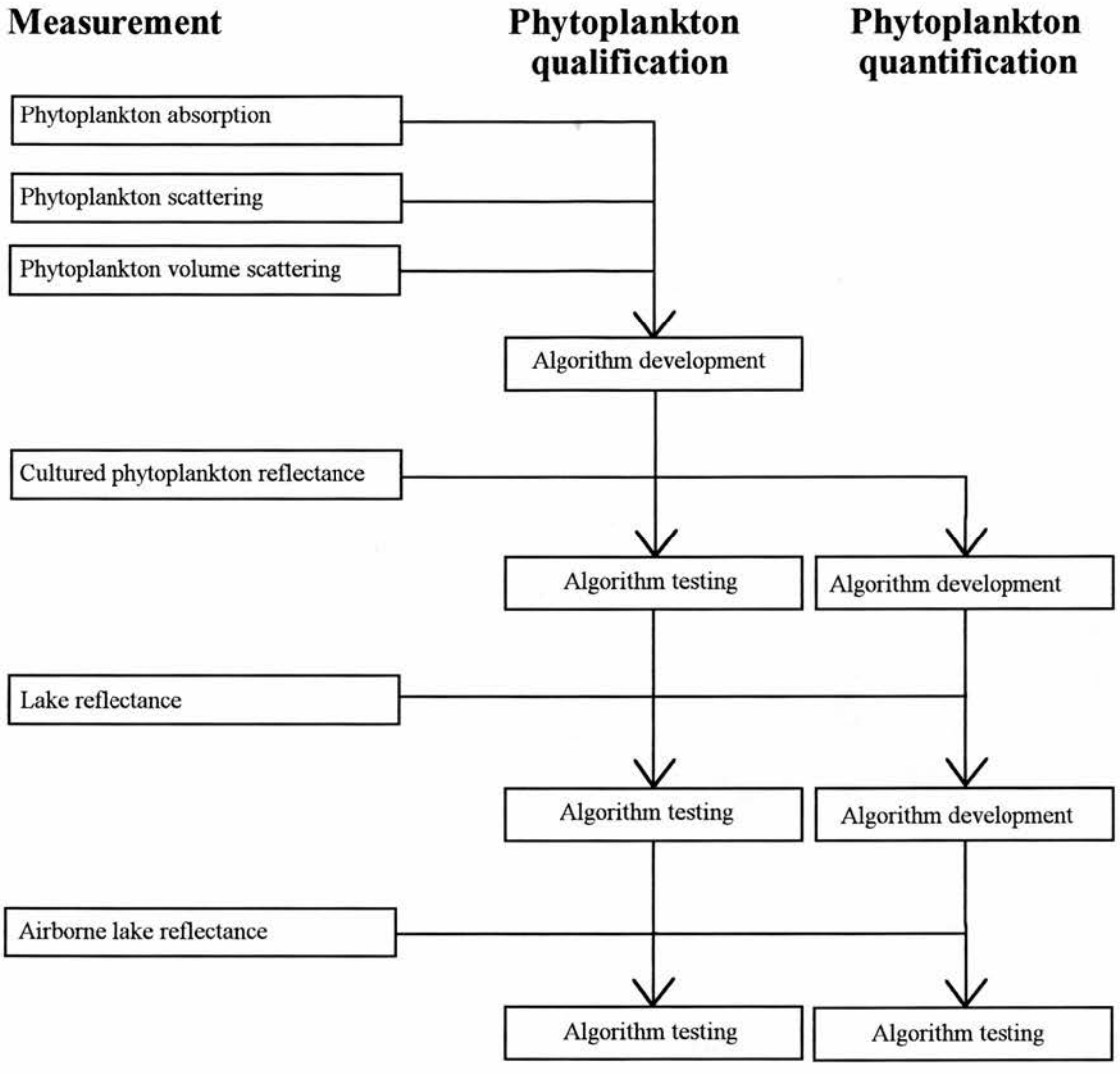

Figure 2. Schematic diagram of the experimental design used in this study, showing how the various optical measurements were used first to develop the quantification and qualification algorithms and then used to test those algorithms.

The thesis chapters follow through this experimental design, describing first the optical property measurements, then the algorithm development, and finally describing the algorithm testing. Chapter Four introduces some of the inherent optical properties of freshwater phytoplankton and considers in detail the pigment composition of the various phytoplankton classes as described in the literature. Pigments which have the potential to act as class marker pigments are identified. Using measurements of $a_{PHY}(\lambda)$ and $b_{PHY}(\lambda)$ the spectral influence of pigment composition is shown to be more significant than the spectral influence of the cell morphology, this is important because within-class pigmentation variation is less significant than within-class cell morphological variation. The strong inverse spectral relationship between the inherent optical property $a_{PHY}(\lambda)$ and the apparent optical property $R_{PHY}(\lambda)$ is highlighted by measurements of pure culture $R_{PHY}(\lambda)$ in a purpose-built laboratory-based experimental tank. The within-class spectral similarities and between-class spectral differences in $R_{PHY}(\lambda)$ are then

demonstrated to be a function of pigment composition. $a_{PHY}(\lambda)$ for controlled mixtures of pure cultures were also made for the purpose of testing the developed class identification routines.

Chapter Five introduces the optical measurements made on Esthwaite Water, a naturally eutrophic Case I water body in the English Lake District where absorption and reflectance spectra with concomitant water quality samples were measured over a two year period. Changes in these spectral measurements are discussed in terms of successional changes in the phytoplankton crop.

The next two chapters deal with the development of the qualitative and quantitative algorithms. Chapter Six develops some new algorithms for the monitoring of phytoplankton crop concentration initially from pure culture $R_{PHY}(\lambda)$ measurements in the experimental tank and then from the Esthwaite Water spectral signatures introduced in Chapter Five. The database of pure culture $a_{PHY}(\lambda)$ introduced in Chapter Four is then used to develop the original phytoplankton class identification routines and algorithms in Chapter Seven. Four main routines are developed: the first is a simple ratio technique which uses observations of spectra at wavelengths where marker pigments absorb, class presence or absence is highlighted by the ratio; the second routine uses canonical discriminant analysis to reduce the dimensionality of the hyperspectral data sets to a set of carefully selected wavelengths; the third routine uses the full hyperspectral data set to calculate the fourth derivative which highlights the subtle pigment features masked in raw spectra thereby comprehensively assessing the pigment composition of the target crop from which an assessment of phytoplankton class can be made; and the forth routine is a neural network application which draws upon the derivative technique. In each case the routines and algorithms, having been developed on pure culture $a_{PHY}(\lambda)$, are tested on the laboratory based measurements of pure culture $R_{PHY}(\lambda)$ made in the experimental tank. In an exception to the main premise used for the development of the identification routines a further qualitative routine was developed directly from the pure culture $R_{PHY}(\lambda)$. This was based upon the number of distinctive peaks present in the visible part of the electromagnetic spectrum. This rather simple approach was limited to distinguishing between the Cyanophytes and non-Cyanophytes.

In Chapter Eight the natural water spectral signatures from Esthwaite Water are used to test the prediction accuracy of the identification routines. The influence of phytoplankton crop type on the quantitative algorithms is also discussed. Conclusions to the whole study, with particular attention being paid to the proposed class identification routines, are drawn in the final chapter.

3.1 Common methods

There are some data processing methods that are common to most of the following chapters. These are discussed below.

3.1.1 Spectral smoothing

In many experiments the true signal amplitudes change rather smoothly, where as noise tends to exhibit rapid, random changes in amplitude within the signal. Smoothing is the removal of noise. Smoothing of the data is often required as a prelude to further treatment of the data (Savitzky and Golay 1964, Turrell 1982) or simply as a 'cosmetic enhancement' to presented data (Enke and Nieman 1976). As long as the underlying signal can be assumed to be smooth, then digital smoothing will not cause significant signal distortion. The spectral smoothing used during this study was the least-squares polynomial fitting routine of Savitzky and Golay (1964) as modified by Turrell (1982) which was converted into a Turbo Pascal v7.0 program (*Saviztky.pas*) (T.J. Malthus pers. comm.).

3.1.2 Spectral splining

To ensure that the ordinate values of measured spectra conformed, certain spectra were splined. If there were too many ordinate values then the spectra were sub-sampled at the appropriate ordinate density. If there were too few ordinate values, or the ordinate values were inappropriately spaced, then the spectra required splining. A Turbo Pascal v7.0 program (*Spline.pas*) was written to fit a first order polynomial to the data set allowing the abscissa value for the user specified ordinate value to be calculated. A first order polynomial was chosen because the spectra typically had reasonably dense ordinate values, so the loss of accuracy was minimal, while the ease of generating the linear fit was high.

3.1.3 Spectral standardisation

To visually enhance spectral differences in absorption or reflectance spectra, without recourse to derivatives, spectra can be standardised. There are a number of potential standardisation methods. Dekker (1993) standardised absorption spectra by calculating the natural-logarithm of the spectra or by calculating the specific absorption coefficient, while reflectance spectra were standardised by choosing an appropriate wavelength value by which to ratio the rest of the data set *e.g.* 550 nm may be used because typically it is the $R_{PHY}(\lambda)$ maximum. Spectra can also be standardised by calculating the equivalent spectra with zero mean and unit variance. This is achieved by using:

$$z = (x - \mu) / \sigma \quad \text{.....3.1.3.1}$$

where z is the new data value, x the original value, μ the mean for the part of the spectra under consideration, and σ is the standard deviation for this region of the spectra.

Chapter Four Optical properties of freshwater phytoplankton

This chapter begins by reviewing the photosynthetic pigments, which through their absorption of light, impart colour to phytoplankton. Phytoplankton absorption is essentially a function of the intra-cellular pigment composition and the pigment composition is a function of class. The pigment compositions of the some classes of freshwater phytoplankton are then considered. To ease comprehension a list of pigment abbreviations is given at the start of this work.

The physical properties of the freshwater phytoplankton species selected for this study are described and measurements of their inherent optical properties are presented. The study concentrated upon pure culture absorption measurements because of the relationship between absorption and reflectance and because of the perceived capacity that phytoplankton absorption spectra have for developing the identification routines, however surrogate measurements for scattering and the volume scattering function are also considered. Variations in the presented phytoplankton absorption spectra are discussed in terms of pigment composition.

Having assessed the inherent optical properties of pure cultures of freshwater phytoplankton, the final section describes a series of measurements of reflectance from pure cultures of phytoplankton made in a purpose built experimental laboratory based tank. As with the phytoplankton absorption spectra variations in phytoplankton reflectance are discussed in terms of the pigment composition.

4.1 Photosynthetic pigments of phytoplankton

Phytoplankton cell size, cell shape, absorption by the cellular organelles, chloroplast arrangement, pigment content and pigment 'packaging' into thylakoids all influence $a_{PHY}(\lambda)$ thereby imposing colour to the cell (Prezlin and Boczar 1986), however it is primarily pigments that colour the cells (Hoepffner and Sathyendranath 1991). Photosynthetic pigments fall into one of three categories: chlorophylls; biliproteins; and carotenoids. The main chlorophylls are CHL*a*, *b*, *c* and *d*; the main biliproteins are PCC, PCE and APC; there are many carotenoids (over 500 occur naturally) including the carotenes and xanthinols (Britton 1983, Rowan 1989). All phytoplankton contain CHL*a*, most contain accessory pigments. CHL*a* is the only pigment necessary for photosynthesis; the other accessory pigments expand the electromagnetic range of the cell absorbing photons at wavelengths other than the Soret and red absorption bands of CHL*a* (Prezlin and Boczar 1986, Owens *et al.* 1987). Successful remote sensing of phytoplankton class will utilise absorption by these pigments to identify the pigment composition of the target bloom which in turn will allow the identification of the phytoplankton class.

The primary purpose of phytoplankton pigments is for photosynthesis, they also have a secondary role in preventing cellular damage by the high energy photons of the near-UV parts of the electromagnetic spectrum. During photosynthesis light energy is harnessed to fix carbon dioxide in a useful form as carbohydrate. This carbohydrate is used as an energy store and as a source of carbon for all other molecules needed by the phytoplankton. Photons at any wavelength from the visible spectrum have enough energy to initiate a photosynthetic response (Britton 1983).

The pigments are contained within a series of discs-like structures called thylakoids which are stacked on top of each other. In Eukaryotic phytoplankton (the Chlorophytes, Bacillariophytes, Cryptophytes and Euglenophytes used in this study are all Eukaryots) thylakoids are confined within a chloroplast; in Prokaryotic phytoplankton (Cyanophytes) the thylakoids exist through-out the cell cytoplasm. There is considerable variation in chloroplast shape and size (Britton 1983). Of the phytoplankton classes considered in this study only the Cyanophytes are Prokaryotic. The intra-cellular location of the pigments is important when considering the package effect (§ 2.1.5.2) because the light harvesting efficiency of photosynthetic pigments is reduced when they are packaged in thylakoids.

Some phytoplankton can adapt their pigment concentrations to suit the prevailing light conditions (Millie *et al.* 1990, Nelson and Prezlin 1990, Ganf *et al.* 1991). Photoadaptation was identified by Ganf *et al.* (1991) as being responsible for the distinct colour differences between samples of *Oscillatoria bourrellyi* taken from the deep-water maximum and from the surface waters of Lake Windermere, England. Vesk and Jeffrey (1977) noted that it is the non-motile phytoplankton, which are unable to regulate their position within the water column, that are most likely to have the ability to photoadapt. Prezlin and Boczar (1986) note that photoadaptation is not restricted to changes in accessory pigment composition with many Chlorophytes photoadapting to low light levels by increasing thylakoid densities within the chloroplast or by increasing chloroplast numbers within the cell. Photoadaptation of pigments raises an issue for this study because inferences based on the optical properties of cultured phytoplankton may not be applicable to phytoplankton from natural water bodies where specific environmental conditions have induced accessory pigment synthesis.

The property of pigments that is most relevant to this work is their ability to absorb light. $a_{PIG}(\lambda)$ is a function of the chemical structure of the pigment molecule. Extracting pigments in solvents causes hypsochromic shifts (decreases in wavelength) or bathochromic shifts (increases in wavelength) of the location of absorption peaks. These shifts are caused by the dissociation of the pigments from their associated proteins and typically range from 6 to 90 nm (Prezlin and Boczar 1986). Hypsochromic shifts prevent the direct comparison of extracted pigment spectra and *in vivo* pigment spectra (Smith and Alberte 1994). This is important to this study because the traditional means for measuring absorption is a spectrophotometer, however such instruments need modification before they can

effectively measure particulate samples (§ 2.1.5.2) and measurements of extracted pigments would be erroneous.

4.1.1 The chlorophylls

The contribution that chlorophylls make to the perceived colour of the plants that contain them is clearly seen in the pervading green of natural vegetation (Britton 1983). Chlorophylls have a basic porphyrin structure with magnesium chelated in the centre. The chlorophylls have several intense absorption bands throughout the visible wavelengths. A particularly intense absorption band occurs at *ca.* 440 nm which is known as the Soret band and is highly characteristic of CHL*a* (Britton 1983). French *et al.* (1967) identified four different types of *in vivo* CHL*a* absorption in the far-red wavelengths, with absorption at 663, 670, 678 and 684 nm. These spectroscopic differences are caused by differences in the intra-cellular microenvironment of the chlorophyll molecules (Britton 1983). The modification of *in vivo* CHL*a* absorption due to the structural orientation of the chlorophyll molecules within the thylakoid membrane has also been noted by Prezlin and Boczar (1986). Prezlin and Boczar (1986) note that some marine Chlorophytes have two to four times as much CHL*b* as CHL*a*, a property which enhances their ability to absorb blue light.

The chlorophylls are unstable compounds and are readily destroyed by light, oxygen, heat, acid and alkali. Even at room temperatures CHL*a* and *b* undergo isomerisation to the closely related phaeopigments CHL*a*' and *b*' which are their immediate breakdown products (Britton 1983, Prezlin and Boczar 1986).

4.1.2 The biliproteins

Biliproteins are acidic, water-soluble, globular proteins with a basic porphyrin structure. The main biliprotiens are PCE, PCC and APC, and the properties of these pigments are dependent on the phytoplankton class, consequently they are given class identifiers *e.g.* Cyanophyte-PCC. The concentration of biliprotein in cells is usually high; values of up to 24% of cellular dry mass have been recorded for PCC. The presence of biliproteins results in characteristic blue and red colours. Quantitatively the two main groups of biliproteins are PCE and PCC. Several types of PCE and PCC can be distinguished spectroscopically, with distinct differences occurring between the absorption spectra of PCE between samples from different phytoplanktonic classes, the same is true of PCC (Britton 1983). Prezlin and Boczar (1986) and Kirk (1994) both make the distinction between Crypto-PCE, Bacillario-PCE and Cyano-PCE with the absorption properties being peculiar to each class. Due to photoadaptation the relative quantities of the different types of biliproteins depends on the spectral quality of the light available to the cells *e.g.* in green light conditions, synthesis of the green-absorbing red pigment PCE is favoured, whereas in red light PCC predominates.

4.1.3 The carotenoids

Almost all carotenoids are derived from tetraterpenes. Those carotenoids containing oxygen are known as xanthophylls. Some 500 naturally occurring carotenoids have been identified (Britton 1983). Carotenoids are known to be involved in a variety of cellular processes including fluorescence suppression (Prezlin and Boczar 1986), photoprotection and light harvesting for photosynthesis which is their most common role (Owens *et al.* 1987).Carotenoids are lipids, they are soluble in organic solvents and can be extracted from natural tissue by polar solvents such as acetone and alcohols. Xanthophylls are virtually insoluble in water. Once isolated, carotenoids are sensitive to heat and light, acids and bases, and oxygen bleaching through contact with air (Britton 1983).

4.1.4 The taxonomic importance of phytoplankton pigments

Taxonomically Phytoplankton are sub-divided into classes, within the classes there are many genera and each genus can exist in several forms of phytoplanktonic species. Taxonomists have identified a number of different classes which can be identified according various criteria. It is apparent from some of the common names used for these classes that pigment content (hence colour) has played an important role in this classification (Christiansen 1962, Faust and Norris 1982, Owens *et al.* 1987) *e.g.* Cyanophytes - Blue-greens and Chlorophytes - Greens.

With the exception of CHL*a* the chlorophylls are distributed among phytoplankton classes in a highly selective manner giving them great chemotaxonomic value, for instance CHL*b* is only present in the Chlorophyceae, and CHLc_1 and c_3 are restricted to Bacillariophyceae, Prymnesiophyceae, Chrysophyceae, Dinoflagellates, and Cryptophyceae. The biliproteins occur in the Rhodophyceae, Cyanophyceae and Cryptophyceae. PCE predominates in Rhodophyceae and PCC in the Cyanophyceae. The primitive Prokaryotic Cyanophyceae produce β-CART and several of its simple derivatives, but many also accumulate MYX. VLX is found in a number of classes where it accounts for only a few percent of total carotenoid (Britton 1983, Prezlin and Boczar 1986, Owens *et al.* 1987, Rowan 1989). These marker pigments are summarised in Table 1.

Table 1. Potential marker pigments for distinguishing between phytoplankton classes as identified in the literature. A bolded entry indicates a class unique pigment.

Phytoplankton class	*Potential class marker pigments*
Cyanophyceae	**C-PCC**, **C-PCE**, **MYX**
Cryptophyceae	CHLc_1, CHLc_3, **R-PCC**, **R-PCE**
Bacillariophyceae	CHLc_1, CHLc_3, **FCX**
Chlorophyceae	**CHL*b***
Euglenophyceae	CHL*b*, **DDX**

Pigments reported as being associated with the five main classes of phytoplankton considered in this study are listed by class in Appendix One along with the author who cited the association. These tables are not intended as a comprehensive list of pigment associations with class. Study of these pigment class associations show that some pigments are largely restricted to one class only. Such

pigments have the potential to act as marker pigments for the purposes of class identification. For example: FCX is a marker pigment for the Bacillariophyceae; DDX for the Euglenophyceae; CHL*b* for the Chlorophyceae and Euglenophyceae; C-PCC and C-PCE are only present in the Cyanophyceae; and R-PCC and R-PCE are present in the Cryptophyceae.

4.1.5 Traditional means of pigment identification

Phytoplankton pigments can be identified using a variety of techniques. CHL*a* is typically assessed spectrophotometrically, as can most pigments for which their extinction coefficients are known (Jeffery and Humphrey 1975). Chromatographic methods, particularly high pressure liquid chromatography (HPLC), provide a more thorough pigment assay, relying upon relative polarity preferences to separate pigments extracted in an apolar substance (Mantoura and Llewellyn 1983).

4.1.6 The pigment database

To assist in the development of the class identification a database of the location of *in vivo* $a_{PIG}(\lambda)$ peaks was compiled from the literature. This pigment database was primarily used as a tool for developing and testing the performance of new analytical techniques for phytoplankton identification. A full listing of the database is presented in Appendix Two. Some of the more salient features highlighted by the database are discussed below.

There is a considerable number of apparent conflicts in the reported *in vivo* absorption features for the different pigments. For instance, the location of the far-red CHL*a* absorption feature is reported to fall between 670 nm (Prezlin and Boczar 1986) and 684 nm (Kirk and Tilney-Basset 1978), and a CHL*b* absorption feature is reported as occurring variously between 640 nm (Prezlin and Boczar 1986) and 658 nm (Smith and Alberte 1994). These disparities may be the result of: differences between cultured and natural phytoplankton populations; different light regimes and light sources used for growing cultured species; differences in measurement methodologies.

Phytoplankton photosynthetic pigments absorb throughout the visible part of the electromagnetic spectrum. Between 401 and 450 nm absorption is dominated by CHL*a*, and by the accessory chlorophylls. Carotenoids dominate from 451 to 550 nm. Absorption by biliproteins dominate from 551 to 650 nm, although accessory chlorophylls also absorb between 601 and 650 nm. The region from 651 to 700 nm is dominated by the far-red CHL*a* absorption maxima at *ca.* 675 nm. There is no region of the visible spectrum that is completely devoid of pigment absorption features. Pigments have different absorption efficiencies and are present in different concentrations within a cell. These two factors combine to give a spectrally complex *in vivo* $a_{PHY}(\lambda)$.

Many of the individual pigments listed in the database have more than one absorption maxima. For instance APC has two identified absorption maxima occurring at 620 and 650 nm. Pigments exhibit

double or even triple absorption bands because of their molecular structure. This property extends the spectral range over-which the cell can absorb.

It is interesting to note the disparity in the reported location of the absorption maxima for the same pigment within the same spectral region. The prime example of this is the absorption by CHL*a* in the far-red, with 14 different reported maxima from 660 to 705 nm although most are within ±4 nm of 676 nm. These reported spectroscopic differences could either be caused by differences in the microenvironment of the chlorophyll molecules (Britton 1983), or by the effects of pigment associations within different photosynthetic systems (Lee 1989), or they may simply reflect the different methods used to determine the peak $a_{PIG}(\lambda)$ wavelength. Full $a_{PIG}(\lambda)$ for the visible spectrum have been published by Bidigare *et al.* (1989) for the pigments CHL*a*, CHL*b*, CHL*c* and PCE.

Richardson and Ambrosia (1997) note that a library of the spectral absorption features of phytoplankton accessory pigments should be developed to facilitate in the identification of these pigments through optical methods. However measuring individual *in vivo* pigment absorption spectra may prove difficult because isolating photosynthetic phytoplanktonic pigments generally requires organic solvents and these tend to induce hypsochromatic shifts on the absorption properties of the pigments.

4.2 Inherent optical properties of phytoplankton

The inherent optical properties of phytoplankton are $a_{PHY}(\lambda)$, $b_{PHY}(\lambda)$, $c_{PHY}(\lambda)$ and the phytoplankton $\beta(\theta)$. As discussed in § 2.2.3 understanding of the inherent optical properties aids in the interpretation of the apparent optical properties. It also assists in the development of class recognition routines and is necessary for both the modelling the underwater light climate and for the development of bio-optical models for remote sensing of phytoplankton biomass (Bricaud and Stramski 1990, Hoepffner and Sathyendranath 1991). Knowledge of phytoplankton absorption $a_{PHY}(\lambda)$, and in particular absorption efficiency, is of ecological importance because it sets constraints on photosynthetic rates and thus determines phytoplankton growth rates. Furthermore $a_{PHY}(\lambda)$ contributes to, and often dominates, the light extinction in lakes thereby limiting the biomass potential of the lake (Agusti 1991). Unfortunately for the study of these inherent properties they all have definitions which rest on a somewhat ideal experimental conditions which are physically unattainable making absolute measurement difficult (Morel and Bricaud 1986).

The ratios $a(\lambda)/c(\lambda)$ and $b(\lambda)/c(\lambda)$ for a particle can be used to determine whether an photon of a particular wavelength is either absorbed or scattered when it encounters that particle. Scattering generates sub-surface irradiance reflectance, while absorption controls the spectral nature of that reflectance (Morel and Bricaud 1986). When $b(\lambda)$ shows little spectral variation, as is the case for

hence the approach that phytoplankton class recognition routines are devised from measurements of $a_{PHY}(\lambda)$ then transposed to measurements of $R_{PHY}(\lambda)$.

This section therefore builds upon the previous section in this Chapter by concentrating on the measurement of the inherent optical property $a_{PHY}(\lambda)$ for pure and mixed cultures. The $a_{PHY}(\lambda)$ database is introduced which is used for the development of class identification routines in Chapter Six. There is also a discussion of $b_{PHY}(\lambda)$ and $\beta(\theta)$.

4.2.1 Selection, description and culturing of the phytoplankton species used in this study

The phytoplankton classification system used here is according to Bold and Wynne (1985). In this study the phytoplankton classes are referred to by their formal class names. Table 2 relates these formal class names to the more commonly used group names.

Table 2. Formal and colloquial phytoplankton class names used in this study.

Formal phytoplankton class name	*Informal phytoplankton class name*
Bacillariophyte or Bacillariophyceae	Diatom (or Brown)
Chlorophyte or Chlorophyceae	Green
Cryptophyte or Cryptophyceae	Cryptomonad
Cyanophyte or Cyanophyceae	Blue-green or Cyanobacteria
Euglenophyte or Euglenophyceae	Euglenoid

In Table 3 the species utilised for this study are described. The species were obtained from the Culture Centre of Algae and Protozoa (CCAP[1]). The particular species were selected on the basis of class membership and ease of culturing. The study concentrated upon species from the classes Chlorophyceae, Cyanophyceae and Bacillariophyceae, all of which are likely to be naturally occurring in the main study lake, Esthwaite Water (Chapter Five). Chlorophytes have a distinctive class marker pigment CHL*b*. Cyanophytes are frequently rich in the biliproteins PCC and PCE although these are not unique to the class and their relative concentrations within a cell can vary through photoadaptation (Prezlin and Boczar 1986). Cyanophytes can also be motile, regulating their position within the water column by adjusting their buoyancy using intra-cellular gas vacuoles. Bacillariophytes are abundant in low light environments. They have siliceous shells which surrounds their vegetative cells. Their cellular morphology, architecture and arrangement of plate structures, are all prime taxonomic indicators of species type (Prezlin and Boczar 1986). Bacillariophytes have a class marker pigment FCX.

[1] Culture Centre of Algae and Protozoa, IFE Ferry House, Far Sawrey, Ambleside, Cumbria, England.

Table 3. Phytoplankton **genus**, *species* and authority listed by class. Also shown are the CCAP strain number and the media type in which the phytoplankton were cultured (JM is Jaworski's Medium, DM is Diatom Media). Also given is a brief physical description of each alga.

Genus *species* Authority	*Class*	*CCAP strain number*	*Culture media*	*Description*
Asterionella *formosa* Hassall 1850	Bacillariophyte	1005/9	DM	Cylindrical cells up to 100 μm long, forming three-dimensional star shape of 4, 8 or 16 cells.
Fragilaria *crotonensis* Kitton 1869	Bacillariophyte	1029/13	DM	Cylindrical cells up to 150 μm long, forming large flat filamentous colonies. Typically a benthic phytoplankton species.
Melosira *varians*	Bacillariophyte	1048/5	DM	Cylindrical cells joined together to produce filaments. Up to 100 μm diameter with numerous brown chloroplasts.
Asterococcus *limneticus* G.M. Smith 1918	Chlorophyte	3/4	JM	Spherical.
Chlamydomonas *reinhardtii* Danegeard 1888	Chlorophyte	11/32b	JM	Large spherical flagellates occasionally with 'eye' spot.
Chlorella *species*	Chlorophyte	211/70	JM	Non-motile unicellular cells which divide into two or four daughter cells.
Dictyosphaerium *pulchellum* Wood 1872	Chlorophyte	222/2d	JM	Non-motile colonial with four small (10 μm diameter) spherical cells connected by thin strands and contained within a mucilage.
Eudorina *elegans* Ehrenberg 1838	Chlorophyte	24/1a	JM	Motile colonies of up to 200 μm diameter formed by small spherical flagellate cells (5 - 15 μm diameter).
Koiella *species*	Chlorophyte	340/1	JM	Up to 200 μm long thin unicells which are non-motile.
Selenastrum *capricornutum* Printz 1913	Chlorophyte	278/4	JM	Small crescent-shaped cells.
Sphaerocystis *schroeterii* Chodat 1897	Chlorophyte	176/6	JM	Spherical cells.
Volvox *aureus* Ehrenberg 1838	Chlorophyte	88/6	JM	Small spherical cells forming large motile spherical colonies which may contain daughter colonies.
Cryptomonas *species*	Cryptophyte	979/66	JM	Oblate cells.
Rhodomonas *lacustris var. nannoplanctica* (Skuja) Javornicky 1976	Cryptophyte	995/3	JM	Oblate, flagellates.
Anabaena *flos-aquae* Brebbisson ex Bornet & Flahault	Cyanophtye	1403/13b	JM	Filamentous, cylindrical cells typically 10 μm diameter with occasional heterocysts cells. Natural populations can be very dense in the summer.
Microcystis *species*	Cyanophyte	1450/13	JM	Small spherical cells forming large colonies contained by mucus.
Oscillatoria *tenuis* Agardhi 1813	Cyanophyte	1459/43	JM	Filamentous cylindrical cells up to 30 μm diameter which can form dense mats.
Euglena *gracilis* Klebs 1883	Euglenophyte	1224/5	JM	A unicellular green flagellate with spindle shaped body up to 100 μm long. Individual chloroplasts are visible.

Phytoplankton were cultured in specially prepared media formulated to ensure the availability of sufficient quantities of the nutrients required for growth. Many of the cultures would have grown faster if they had been grown in an organic based media but the use of organic media has the draw-back of being very susceptible to contamination by bacteria. The risk of a culture becoming non-axenic (contaminated) is reduced by using non-organic media. Two inorganic media formulations were required for the culturing of the species included in this study. Table 3 shows which media were used for the culturing of the different species. Principally, Diatom Media (DM media, Appendix Three Beakes, Canter and Jaworski 1988) was used for the Bacillariophytes which require silicates for growth. The other species were all cultured in the Jarworski's Blue-green Medium (Appendix Three).

In the sterile environment of a laminar flow cabinet the innoculum of each species was added to a volume of prepared media in a spherical glass flask. The flasks were sealed with a sponge bung and placed in a culturing cabinet where they were incubated at a constant 19°C and received light at 12 hour intervals from four 30 W 'Cool white' fluorescent tubes. The flasks were shaken daily to ensure the media were homogenous as some of the species had a tendency to accumulate through flocculation or settling. The 'colour' of the cultures was regularly noted to ensure that the culture remained in the exponential phase of growth. While in this phase the cells can be assumed to be healthy and cell detritus from the death of cells to be minimal. The exponential growth phase was confirmed by cell counts at weekly intervals.

4.2.3 Phytoplankton absorption

The factors governing $a_{PHY}(\lambda)$ were discussed in § 2.1.5.2.1. Equation 2.1.5.2.1 shows that $a_{PHY}(\lambda)$ is essentially a function of pigment composition although other parts of the cell may have spectral absorption properties. Furthermore cell morphology can influence $a_{PHY}(\lambda)$ through the particle effect. $a_{PHY}(\lambda)$ by natural populations of phytoplankton is difficult to study; it can not be assumed that naturally occurring phytoplankton species have identical $a_{PHY}(\lambda)$ to cultured species. Moreover absorption in cultures can be altered by the formation of blooms, the disaggregation of flocs or by contaminating bacteria on the surface of cells in non-axenic cultures (Morel *et al.* 1993). This not-with-standing the measurement of the optical properties of cultured phytoplankton is useful for the development of class identification routines.

As discussed in § 2.1.5.2 $a_{PHY}(\lambda)$ is typically measured in a spectrophotometer. Spectrophotometers pass a collimated, monochromatic beam simultaneously through two cuvettes, one containing the sample medium, the other a reference medium. The sample's absorbance $D(\lambda)$ is the difference between the two transmitted beams. $D(\lambda)$ is unit-less but is converted to absorption $a(\lambda)$ m^{-1}, by:

$$a(\lambda) = 2.303\left(\frac{D(\lambda)}{r}\right) \quad \text{.....4.2.3.1}$$

where r is the pathlength (m) of the sample cuvette, and 2.303 converts from $\log_{10}$ to natural log.

The measurement of a dissolved substance presents little difficulty to traditional spectrophotometers however when the sample possesses significant scattering properties, as is the case for phytoplankton, then those photons scattered outside the receiving slits of the photomultiplier are erroneously recorded as absorbance leading to the measurement of apparent absorbance $\chi(\lambda)$ (Davis-Colley *et al.* 1986). Modifications to the traditional spectrophotometer configuration are necessary for true absorbance of a particulate sample to be measured.

Over the study period two different absorption methods were employed using three different spectrophotometer set-ups. The two methods were those of Davis-Colley *et al.* (1986) and those of Bricaud and Stramsky (1990) both of which are elaborated below. All the $a_{PHY}(\lambda)$ were measured at 1 or 2 nm intervals from 380 to 800 nm (the spectra at 1 nm intervals were splined to 2 nm intervals to maintain consistency of the over-all data set according to the method in § 3.1.2). Any data smoothing was conducted according to the routine described in § 3.1. The measurements of $a_{PHY}(\lambda)$ were pooled to form the pure culture and mixed-culture $a_{PHY}(\lambda)$ databases.

4.2.3.1 The Davis-Colley *et al.* (1986) method for measuring phytoplankton absorption

Davis-Colley *et al.* (1986) note that for a particulate sample measured absorption in a traditional spectrophotometer set-up is actually apparent absorption $\chi(\lambda)$ which is greater than true absorption $a(\lambda)$ by the angular distribution and intensity of $b(\lambda)$:

$$\chi(\lambda) = a(\lambda) + (1 - \varepsilon)b(\lambda) \quad \text{.....4.2.3.1.1}$$

where ε is the fraction of light scattered but still collected by the spectrophotometer's photomultiplier. $\chi(\lambda)$ is therefore a function of the geometry of the spectrophotometer sample chamber and will increase as the cuvette is moved away from the entrance slits to the photomultiplier because more scattered light falls outside these entrance slits. In normal spectrophotometers ε is generally greater than 50%. Fitting an integrating sphere increases ε to more than 90%. Coefficients measured when ε is high are denoted with an asterix. With an integrating sphere the $(1 - \varepsilon^*)b(\lambda)$ part of Equation 4.2.3.1.1 is very small:

$$\chi^*(\lambda) = a(\lambda) + (1 - \varepsilon^*)b(\lambda) \approx a(\lambda) \quad \text{.....4.2.3.1.2}$$

hence $\chi^*(\lambda)$ has no wavelength dependent scattering component. The relative scattering coefficient, $\Delta\chi(\lambda)$ is the difference between $\chi(\lambda)$ and $\chi^*(\lambda)$ and can be used to indicate the wavelength dependency of scattering:

$$\Delta\chi(\lambda) = \chi(\lambda) - \chi(\lambda)^* = (\varepsilon^* - \varepsilon)\, b(\lambda) \quad \text{.....4.2.3.1.3}$$

If $\varepsilon^* - \varepsilon$ is regarded as a constant for a particular spectrophotometer set-up then $\Delta\chi(\lambda)$ is proportional to $b(\lambda)$. However $\varepsilon^* - \varepsilon$ is difficult to quantify therefore measurement of $\chi(\lambda)$ and $\chi^*(\lambda)$ does not readily provide an estimate of $b(\lambda)$. Nevertheless $\Delta\chi(\lambda)$ still indicates the wavelength trend in scattering and can be used to obtain $a(\lambda)$. Making one measurement when $(1 - \varepsilon)b(\lambda)$ is small (χ^*) and

one when $(1 - \varepsilon)b(\lambda)$ is large (χ) allows the calculation of $\Delta\chi(\lambda)$ from Equation 4.2.3.1.3. $\Delta\chi(\lambda)$ then needs correction for residual scatter which can be performed using a wavelength where $a_{PHY}(\lambda)$ is negligible. Most workers believe that beyond 700 nm a_{PHY} is low, and by 750 nm there is no a_{PHY} (Duysens (1956), Bannister (1988) and Sathyendranath *et al.* (1987)) so at 750 nm all χ^* can be regarded as residual scatter:

$$\chi^*(750) = (1 - \varepsilon^*)b(750) \quad \text{.....4.2.3.1.4}$$

Thus with the two measurements χ and χ^* and the correction for residual scatter $a(\lambda)$ can be estimated according to:

$$a(\lambda) = \chi^*(\lambda) - \chi^*(750)\left(\frac{\Delta\chi(\lambda)}{\Delta\chi(750)}\right) \quad \text{.....4.2.3.1.5}$$

The Davis-Colley *et al.* (1986) method described above was modified for two different spectrophotometer set-ups: a Hitachi U2000 and a Perkin Elmer Lambda 16. At IFE a Hitachi U2000 was used to make the measurements of $a_{PHY}(\lambda)$ and for the measurements of $a_{SES}(\lambda)$ for the Esthwaite Water sampling described in Chapter Six. The Hitachi U2000 has a high resolution concave diffraction grating achieving a spectral bandpass of 2 nm. It has a wavelength accuracy of $\pm$ 0.4 nm which allows a photometric accuracy of $\pm$ 0.002 absorbance units. The Hitachi U2000 was operated in 'remote' mode which allowed downloading of the spectra via an RS232 connector and 'Kermit' v3.0 data-transfer software. The Hitachi U2000 was calibrated by IFE technicians on a quarterly basis using potassium dichromate solution to check photometric accuracy, and holmium and didymium filters to check wavelength accuracy.

There was no available integrating sphere attachment for the Hitachi U2000, thus the χ^* measurement could not be conducted in the same manner as Davis-Colley *et al.* (1986). To provide a similar measurement with a significantly reduced $(1 - \varepsilon)b$ value the spectrophotometer was modified to allow the cuvette holders in the sample chamber to be placed adjacent to the entrance slits of the photomultiplier thereby increasing the acceptance angle of the photomultipliers. Diffusing glass plates were then positioned between the sample and reference cuvettes and the entrance slits to their respective photomultipliers. These plates increase the fraction of the light scattered, but not absorbed, by the phytoplankton cells which reaches the photomultiplier (Yentsch 1962, Haardt and Maske 1987, Bannister 1988). Positioning the cuvettes adjacent to the entrance to the photomultiplier and using the diffusing glass plates allow the minimisation of $(1 - \varepsilon)b(\lambda)$ so the measurement can be treated as $\chi^*(\lambda)$. Measurements were then made with the cuvettes placed 15 cm from the photomultiplier entrance slits thus maximising $(1 - \varepsilon)b$ to give $\chi(\lambda)$. $\chi(\lambda)$ and $\chi^*(\lambda)$ could then be applied to Equation 4.2.3.1.5 to derive true $a_{PHY}(\lambda)$.

All measurements were made at 1 nm intervals from 380 to 800 nm with the scan speed of 400 nm min^{-1}. The reasonably fast scan speed was required to reduce the affect of settling upon the sample's optical properties. 10 mm optically matched glass cuvettes were used for all measurements. Prior to all

measurements, and following every change in the spectrophotometer set-up, a base-line measurement was performed on distilled water. Both the sample and reference spectra were automatically adjusted for this base-line. The reference cuvette always contained the same media used to suspend the phytoplankton hence the absorption properties of this suspending media were automatically removed from the measurements. All measurements were repeated three times and the average taken. Prior to each spectra the sample in the cuvette was shaken to prevent excessive settling of particles over the scan period.

It is believed that this method accurately measures the spectral shape of the $a_{PHY}(\lambda)$ through the visible and near-IR. However, for a number of reasons it is unlikely that the method resulted in a precise measurement of the magnitude of the absorption coefficient:

- Beer's Law does not hold in the presence of multiple scattering because the photon pathlength is artificially increased. In order to achieve an acceptable signal-to-noise ratio in the presence of the diffusing plates the phytoplankton concentrations had to be high thereby increasing the likelihood of multiple scatter events. Haardt and Maske (1987) measured a phytoplankton dilution series from 75 to 300 μg CHL*a* L^{-1} and found no change in the specific absorption coefficient (Equation 2.1.4.2.2) suggesting that multiple scattering was not present in their data set. Similarly Sathyendranath *et al.* (1987) found that Beer's Law held for concentrations as high as 10^3 μg CHL*a* L^{-1}, the concentrations of the phytoplankton cultures measured during this study were typically less than this;
- placing a diffusing glass plate between the cuvettes and the photomultiplier would cause some reflectance at the interface, and this reflectance increases for the beam which has passed through a suspensoid sample because it is less collimnated. Thus the χ^* measurement made with the diffusing glass plate will be higher than it would otherwise be leading to an over estimation of $a_{PHY}(\lambda)$;
- the amount of backscattering within the sample was not known. Backscattered photons would be not be recorded by the photomultiplier, nor would they be accommodated for by Equation 4.2.3.1.5. Unaccountable backscattering would cause an over estimation of $a_{PHY}(\lambda)$;
- similarly the amount of side scattered photons lost through the cuvette side walls was an unknown;
- correcting for residual scattering using $\chi(750)$ according to Davis-Colley *et al.* (1986) assumes there is no absorption at this wavelength. Measurements of the absorption spectra for pigments extracted in organic solvents tend to confirm this assumption, however Bricaud *et al.* (1983) suggest that apparent absorption at this wavelength may be due to non-extractable pigments. If this assumption is incorrect then the magnitude of $a_{PHY}(\lambda)$ will be underestimated.

At Delft University[2] a Perkin-Elmer Lambda 16 spectrophotometer fitted with an integrating sphere attachment running operating system version 4.01 was available thereby allowing the Davis-Colley *et al.* (1986) method to be more precisely followed. The particulars of this set-up are fully described in Hakvoort (1994). Operating in transmission mode, using culture media as a reference, the spectra of samples in 10 mm cuvettes were measured from 400 to 800 nm at 1 nm intervals. The degree of particle settling was checked before each sample was measured by monitoring for a reduction in absorbance at a single wavelength for sixty seconds. If settling was apparent then a micro-bead stirrer was used to maintain an homogenous suspension, if no settling was apparent the stirrer was not used because it introduced noise through turbulence. All spectra were measured in triplicate and the averaged used.

4.2.3.2 The Bricaud and Stramski (1990) filter pad method for measuring $a_{PHY}(\lambda)$

At the Centre for Limnology[3] the filter pad method first introduced by Shibata (1958) and recently modified by Bricaud and Stramski (1990) was used to measure $a_{PHY}(\lambda)$. The phytoplankton were filtered onto 100 mm diameter Whatman GF/F filter. The absorbance of the cells on the filter $Df_{PHY}(\lambda)$ was measured directly in a Phillips PU 8800 UV/VIS spectrophotometer referenced against a blank wetted filter. Both filters were placed immediately in front of the photomultiplier maximising the acceptance angle. Absorption coefficients were calculated according to Equation 4.2.3.1 using the optical pathlength z (where z = V/S and V is the volume of water filtered in litres, and S is the filtering area of the filter in cm^2). To account for the optical effects of measuring the phytoplankton on a filter pad, the absorbance by the sample suspension was then measured using distilled water as a reference in 10 mm optically matched cuvettes placed adjacent to the entrance slits of the photomultiplier. It was assumed that there should be no spectral difference between absorption in suspension and absorption on the filterpad but there should be magnitudinal differences attributable to the absorption by the filter pad itself. Comparison of the filter pad and cuvette absorption spectra allows the determination of the pathlength amplification factor B which arises from the removal of the cellular particles from suspension and their compaction upon the surface of the filter pad. The magnitude of B will vary with filter paper thickness, spectrophotometer and sample and must therefore be calculated on every occasion according to:

$$B = 1.63\left(a(\lambda) - a(750)\right)^{-0.22} \quad \text{.....4.2.3.2.1}$$

(Bricaud and Stramski 1990). The a(750) measurement corrects for scattering by assuming that $\chi(750)$ is entirely due to scatter. The absorption of the seston on the filterpad is calculated according to:

$$a_{(\lambda)} = \frac{\left(2.303\left(a(\lambda) - a(750)\right)d\pi\right)}{(VB)} \quad \text{.....4.2.3.2.2}$$

---

[2] Delft University of Technology, Delft, The Netherlands.
[3] Netherlands Institute of Ecology, Centre for Limnology, Rijksstraatweg 6, 3631 AC Nieuwersluis, The Netherlands.

where d is the diameter of the filter pad (mm) and V is the volume of filtered sample.

4.2.4 The pure culture phytoplankton absorption spectra

In all, 88 pure culture $a_{PHY}(\lambda)$ were measured on 18 species from 5 classes using the two methods described above. The $a_{PHY}(\lambda)$ were all measured when the phytoplankton were in exponential growth. Phytoplankton CHL*a* concentration was measured using the hot methanol extraction method of Talling and Driver (1963) which is more fully discussed in § 6.1. The specific absorption spectra $a_{PHY}^*(\lambda)$ was then calculated according to Equation 2.1.5.2.2 because $a_{PHY}^*(\lambda)$ is useful for comparison of measurements of $a_{PHY}(\lambda)$ performed at different biomass concentrations.

The $a_{PHY}(\lambda)$ measured in this study by the various methods were similar to those previously measured by (Davis-Colley *et al.* 1986, Hoepffner and Sathyendranath 1993, Roesler *et al.* 1989). All the phytoplankton species have broadly similar $a_{PHY}^*(\lambda)$ due to the absorption properties of the ubiquitous CHL*a*. Below 400 nm the $a_{PHY}^*(\lambda)$ is low, it rises from 400 to the Soret band absorption at 440 nm where absorption by CHL*a* is strongest, there are some absorption features between 440 and 500 nm but typically absorption falls to a minima at 550 nm, slowly rising to 650 nm then rising more steeply to 675 nm which is the far-red CHL*a* absorption. Absorption then falls quickly beyond 700 nm. Theoretically at 675 nm it would be expected that a_{PHY}^* would equate for the various species (Bricaud *et al.* 1983), but this is typically not the case. Equal $a_{PHY}^*(675)$ values assume that the CHL*a* for each species has identical absorption characteristics and that extraction of CHL*a* by hot methanol is consistent for each species allowing accurate spectrophotometric determination of *in vivo* CHL*a*. Both these assumptions are unlikely to hold. Furthermore the influence of the particle effect (§ 2.1.5.2) must also be removed before $a_{PHY}^*(675)$ will equate for the various species.

Sample $a_{PHY}^*(\lambda)$ from the various classes are presented below. The between-class spectral differences are discussed in terms of the class pigment compositions. Pigments are identified from their reported absorption features (§ 4.1 and Appendix Two) however it should be noted that these pigment identifications are not definitive.

4.2.4.1 Chlorophyte specific absorption spectra

A representative sample spectra from the 34 Chlorophyceae $a_{PHY}^*(\lambda)$ are shown in Figure 3. The Chlorophytes have a broad $a_{PHY}^*(\lambda)$ shoulder from 450 to 500 nm, both *Chlamydomonas* and *Scenedesmus* have distinct peaks at 470 nm. From the 550 nm minima $a_{PHY}^*(\lambda)$ rises to 675 nm with shoulders at 575, 630 and 650 nm. These features give the far-red CHL*a* peak an asymmetric appearance skewed to the longer wavelengths. These $a_{PHY}^*(\lambda)$ features are caused by CHL*b* (470 and 650 nm), PCE (575 nm) and PCC (630 nm). These biliproteins have not previously been identified in Chlorophytes however there are no other pigments reported as absorbing between 570 and 630 nm.

The skewed far-red absorption peak is caused by the juxtaposition of absorption by CHL*b* (650 nm) and CHL*a* (675 nm).

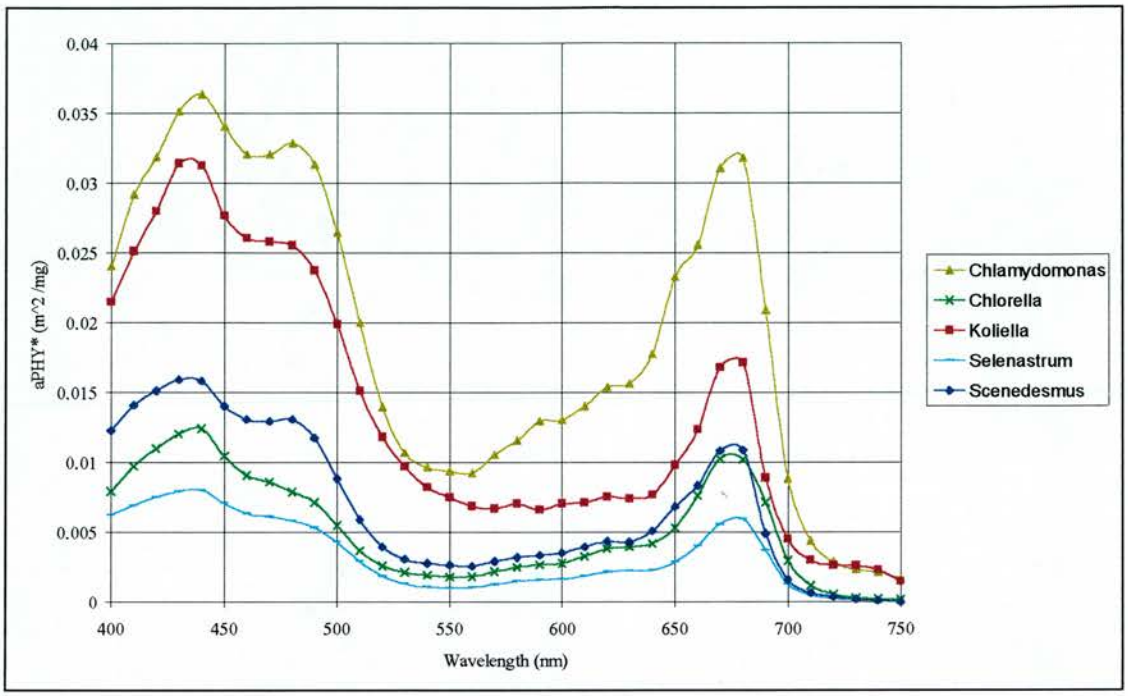

Figure 3. Spectra from pure culture $a_{PHY}^*(\lambda)$ for five species from the class Chlorophyceae.

These spectra compare well to published Chlorophyte $a_{PHY}^*(\lambda)$. Prezlin and Boczar (1986) publish the $a_{PHY}^*(\lambda)$ for a *Chlorella* which has peaks at 440, 650 and 678 nm, a trough at 540 nm and shoulders at 470, 590 and 620 nm. Davis-Colley *et al.* (1986) publish an $a_{PHY}^*(\lambda)$ for a *Scenedesmus* and a *Chlamydomonas*. The comparison between the *Scenedesmus* $a_{PHY}^*(\lambda)$ presented here and the publish spectra is good with the slight exception of the published spectra's trough occurring at 520 nm which is slightly longer than the presented $a_{PHY}^*(\lambda)$. There is broad similarity between the two *Chlamydonmonas* $a_{PHY}^*(\lambda)$, however between the 550 nm trough the published spectra exhibits more shoulder features before the 675 nm peak than does the spectra presented here.

4.2.4.2 Cyanophyte specific absorption spectra

Representative spectra from the 27 Cyanophyte $a_{PHY}^*(\lambda)$ are shown in Figure 4. The major $a_{PHY}(\lambda)$ associated with the Cyanophytes is the third absorption peak centred on 630 nm which is of similar magnitude as the far-red CHL*a* absorption peak. There are also shoulders at 490 and 580 nm. The far-red CHL*a* absorption peak is less skewed than for the Chlorophytes. The 630 nm absorption is caused by PCE and the 580 nm shoulder by the associated biliprotein PCC. The 490 nm shoulder is probably a xanthinol absorption feature.

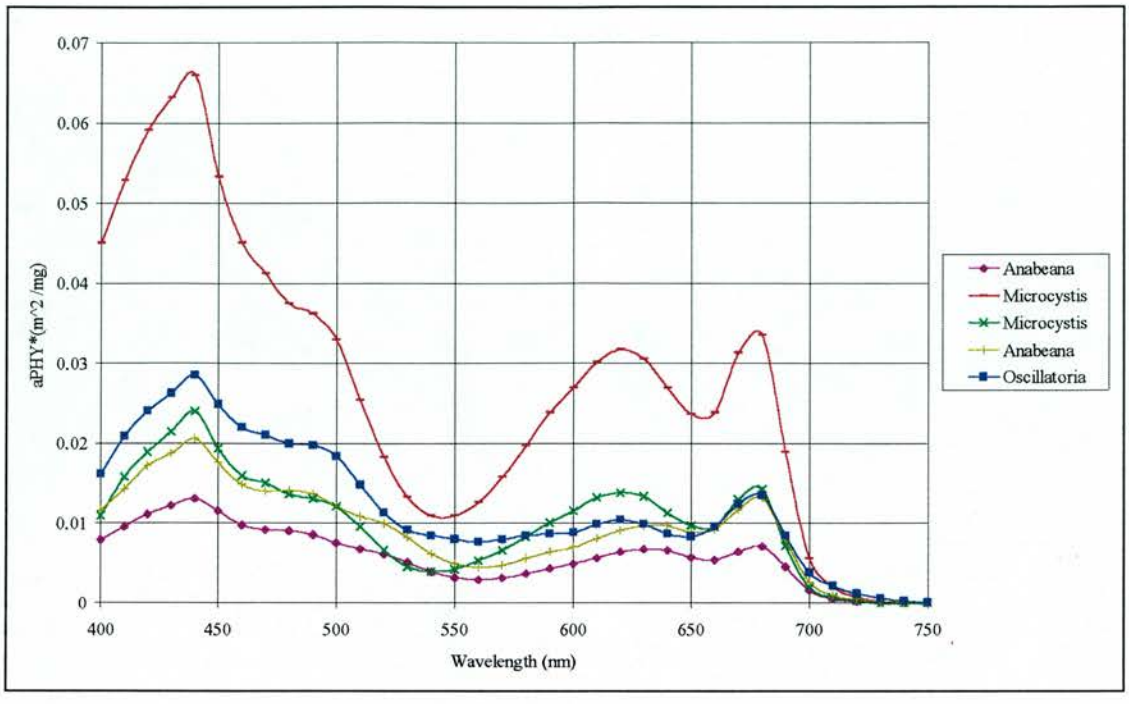

Figure 4. Spectra of pure culture $a_{PHY}^*(\lambda)$ for three species from the class Cyanophyceae.

Again the similarities between published Cyanophyte $a_{PHY}^*(\lambda)$ and those presented above are strong. Davis-Colley *et al.* (1986) present the $a_{PHY}^*(\lambda)$ for an *Anabaena* species which differs slightly from the *Anabaena* spectra present here as it troughs at a shorter wavelength (540 nm compared to 560 nm).

4.2.4.3 Bacillariophyte specific absorption spectra

Representative spectra from the 17 Bacillariophyte $a_{PHY}^*(\lambda)$ are shown in Figure 5. The Bacillariophytes exhibit more within-species spectral variation than the Chlorophytes and Cyanophytes. There is an absorption shoulder at 480 nm for the *Fragilaria* spectra and a peak at 490 nm for the *Melosira* spectra both of which are absent from the two *Asterionella* spectra. The *Melosira* and two *Asterionella* spectra show a shoulder at 525 nm which is absent from the *Fragilaria* spectra. All spectra show a small 630 nm peak, and the *Fragilaria* spectra shows a small 650 nm peak. The *ca.* 490 nm peak is likely to be one of a range of xanthinols (FCX, DDX, DTX, VLX or ZAX) however the case for it being FCX is enhanced by the presence of the 525 nm feature which can only be caused by FCX. The 630 nm peak is due to the various forms of CHL*c*, and the 650 nm *Fragilaria* absorption feature may be due to a previously unreported CHL*c* absorption peak. Davis-Colley *et al.* (1986) present the $a_{PHY}^*(\lambda)$ for a Bacillariophyte (*Navicula*) which generally compares well with those Bacillariophyte $a_{PHY}^*(\lambda)$ presented here.

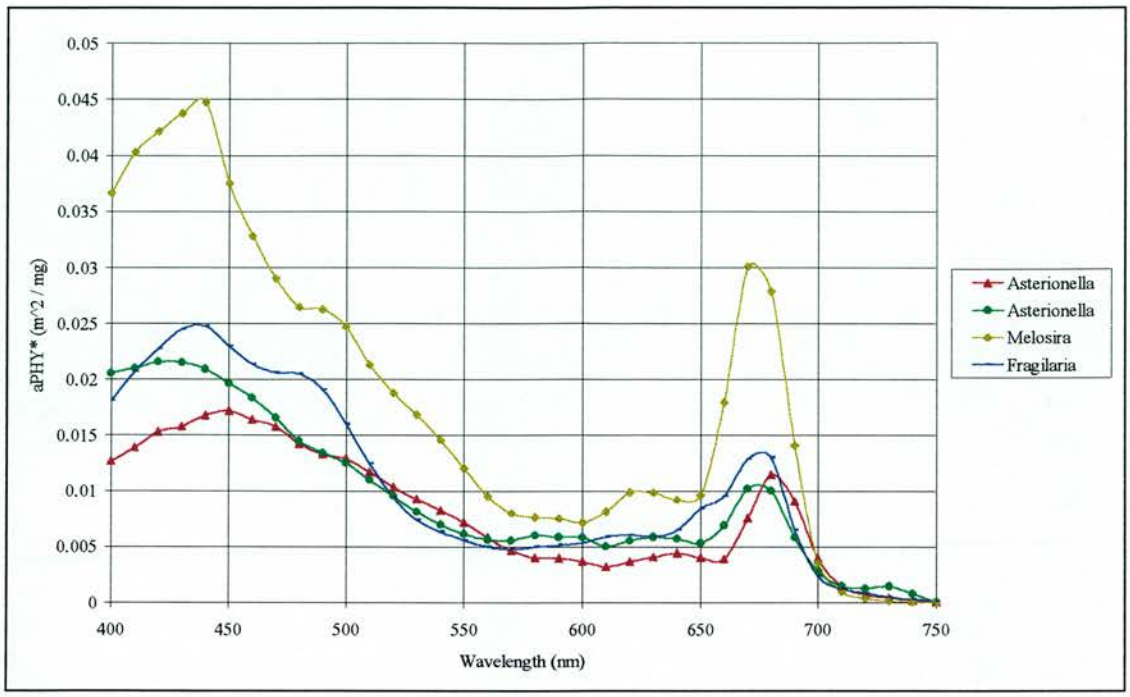

Figure 5. Spectra of pure culture $a_{PHY}^*(\lambda)$ for three species from the class Bacillariophyceae.

4.2.4.4 Cryptophyte and Euglenophyte specific absorption spectra

Representative spectra from the eight Cryptophyte and six Euglenophyte $a_{PHY}(\lambda)$ are shown in Figure 6. The Cryptophyte $a_{PHY}^*(\lambda)$ are the most spectrally complex of the five classes, they also show the most within-class differences. The absorption features common to both the *Rhodomonas* and *Cryptomonas* $a_{PHY}^*(\lambda)$ are at 470 and 500 nm, the *Cryptomonas* $a_{PHY}^*(\lambda)$ also has a feature at 570 nm and the *Rhodomonas* has features at 550, 600, 620 and 645 nm. The 470 nm feature is probably due to the presence of CHL*c*, the 500 nm feature is due to FCX, the 550 nm feature is probably due to either PCE or PCC, the 570, 600 and 620 nm features are due to the presence of PCE while the 645 nm feature is caused by PCC.

The two *Euglena* $a_{PHY}^*(\lambda)$ are very similar with shoulders at 490, 580, 620 and 655 nm. These are due to DDX; possibly PCE; possibly PCE, PCC or APC; and CHL*b* respectively. The presence of biliprotiens in Euglenophytes is not confirmed by the references in the literature, however there are no other pigments reported as absorbing between 570 and 630 nm.

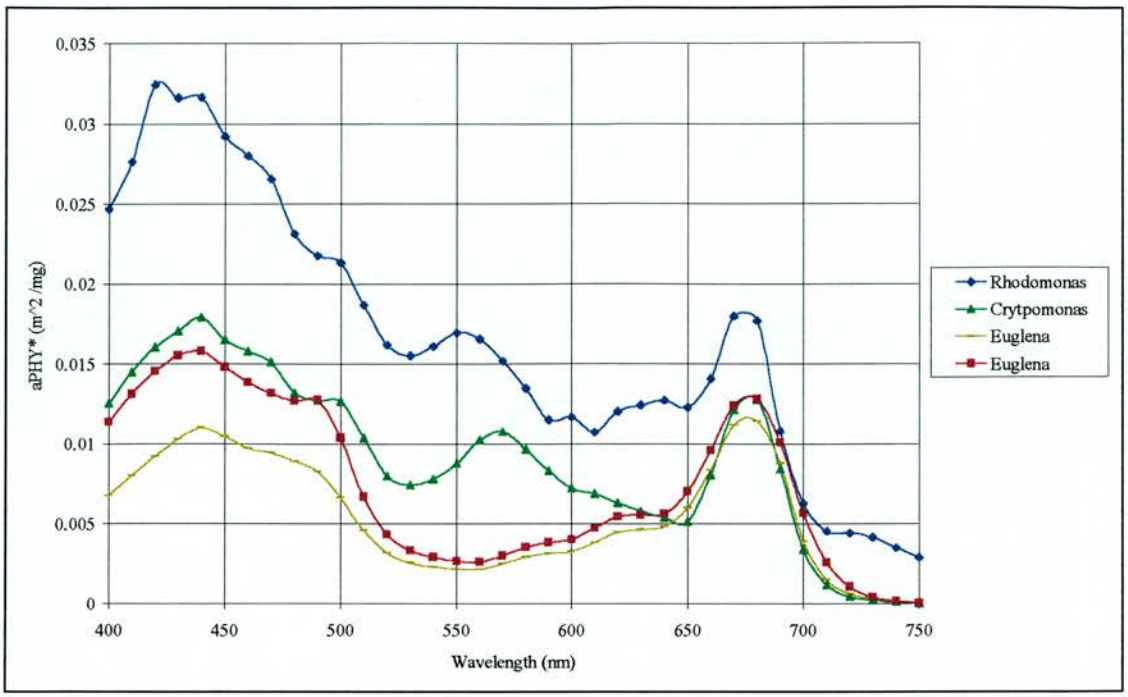

Figure 6. Spectra for pure culture $a_{PHY}^*(\lambda)$ for two species from the class Cryptophyceae (*Rhodomonas* and *Cryptomonas*) and one species from the class Euglenophyceae.

4.2.4.5 Between-class comparisons of specific absorption spectra

The species variation in pure culture $a_{PHY}^*(\lambda)$ as measured during this study are similar to those of Bricaud *et al.* (1988) who reported variations in $a_{PHY}^*(435)$ from 0.02 to 0.075 m^2mg^{-1}. The between-class similarities in $a_{PHY}^*(\lambda)$ are more obvious than the between-class differences due to the dominant absorption properties of the ubiquitous CHL*a*. However there are perceptible between-class differences and within-class similarities which concur with the findings of Davis-Colley *et al.* (1986). The Cyanophytes have the most unique spectral form with the strong absorption peak at 630 nm caused by PCE. The Cryptophytes have an interesting absorption peak at 560 nm also caused by PCE. That the same pigment can have two absorption centres is not unusual with the absence of strong PCE absorption at 630 nm in the Cryptophytes and the absence of strong PCE absorption at 560 nm in the Cyanophytes highlighting the tendency for biliprotiens to have different absorption properties depending upon their class association as was discussed in § 4.1.2. Finally the CHL*b* absorption features distinguish the Chlorophytes from the other classes with the exception of the Euglenophytes.

4.2.5 Specific absorption spectra for mixed cultures of phytoplankton

The $a_{PHY}(\lambda)$ for 50 controlled mixtures of pure cultures in exponential growth were also measured. These measurements were performed for the testing of the class identification algorithms and are described here. For each mixture the relative proportions of the two classes in the mixture were ascertained through prior measurement of the CHL*a* concentrations of the pure cultures. In some instances the two species in the mixture were drawn from the same class. For reasons of clarity only a sample of these mixed culture $a_{PHY}^*(\lambda)$ are shown in Figure 7. The $a_{PHY}^*(\lambda)$ for the mixed cultures are

typically more complex than for the pure cultures because of the increased complexity of the pigment composition and its ultimate influence on the $a_{PHY}(\lambda)$.

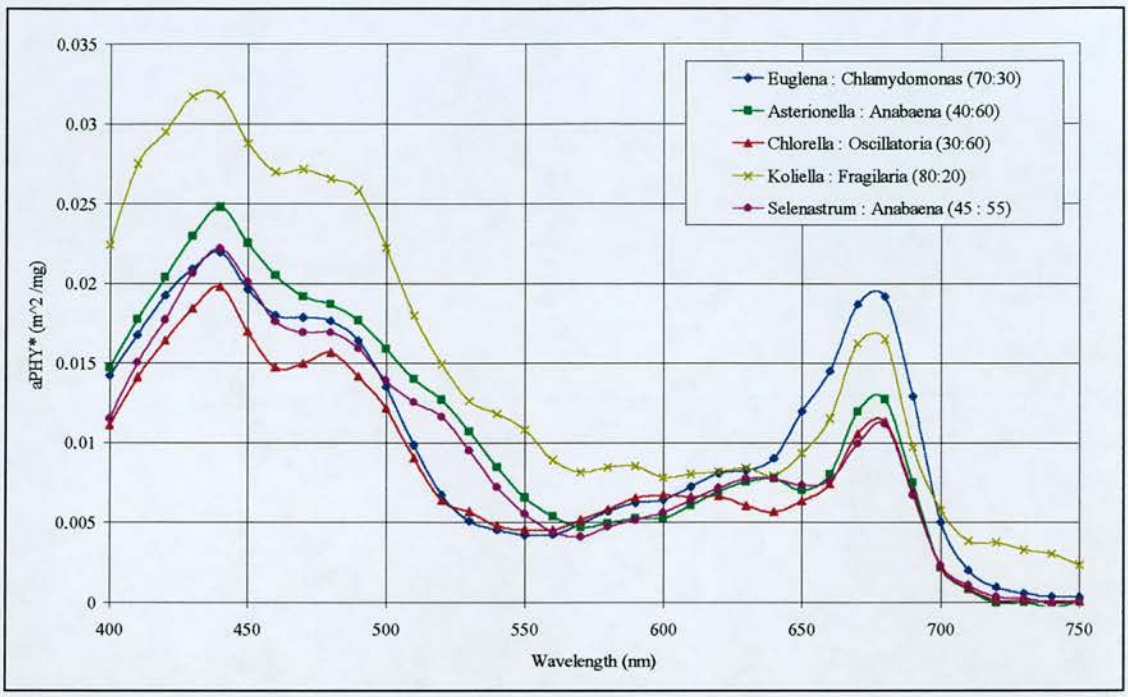

Figure 7. Spectra for five examples of the mixed culture $a_{PHY}^*(\lambda)$. The legend shows the relative percentage of each class in the mixture.

4.2.6 Phytoplankton scattering

The influence of scattering $b(\lambda)$ upon reflectance has been discussed in Chapter Two. Much of the analysis in this study makes the assumption that the majority of the spectral variation in $R_{PHY}(\lambda)$ in the visible spectrum is caused by the spectral variation in $a_{PHY}(\lambda)$ induced by between-class differences in pigment content. $b_{PHY}(\lambda)$ is assumed to play little role in determining the spectral variation in $R_{PHY}(\lambda)$ (§ 2..1.5.2). Making this assumption allows the study to concentrate on the pigment induced variations in $R_{PHY}(\lambda)$ ignoring the spectral implications of cell morphological differences. This is necessary because there is little within class similarity in cellular morphology, while the within-class similarity in pigment composition has already been highlighted.

$b_{PHY}(\lambda)$ is influenced by absorption through anomalous dispersion, by cell size and by the cellular refractive index (§ 2.1.5.2). Davis-Colley *et al.* (1986) note that $b_{PHY}(\lambda)$ is typically an order of magnitude higher than the $a_{PHY}^*(\lambda)$ and hence it is the most important contributor to light attenuation by phytoplankton $c_{PHY}(\lambda)$. The magnitude of $b_{PHY}(\lambda)$ varies greatly between species. Bricaud *et al.* (1983) found variations in $b_{PHY}(550)$ to range from 0.08 to 0.62 m^2 mg CHLa^{-1}, consequently the magnitude of backward scattering will also be strongly species dependent.

$b_{PHY}(\lambda)$ and its directional components (backward scattering $b_{bPHY}(\lambda)$ and forward scattering $b_{fPHY}(\lambda)$) are not easily measurable, furthermore direct measurement of $b_{PHY}(\lambda)$ is confounded by fluorescence emissions from pigments (Morel *et al.* 1993). Indirect measurements of $b_{PHY}(\lambda)$ depends upon measurements of $a_{PHY}(\lambda)$ and $c_{PHY}(\lambda)$ and calculation of the difference between the two (Equation 2.1.7.3). Theoretical calculations of $b_{PHY}(\lambda)$ rely upon Mie theory and a knowledge of the cell size, shape and refractive indices for each species and there are many errors which compound such estimates (Morel and Bricaud 1986).

The Davis-Colley *et al.* (1986) method for measuring absorption of a particulate sample involves calculation of the unitless relative scattering spectra $\Delta\chi(\lambda)$ (§ 4.2.3.1). They have shown that this can be used to compare $b_{PHY}(\lambda)$ for different classes. Representative spectra for the pure culture $\Delta\chi(\lambda)$ are presented below. Figure 8 to Figure 10 show a sample of the specific relative scattering spectra $\Delta\chi^*(\lambda)$. The Chlorophytes $\Delta\chi^*(\lambda)$ (Figure 8) varied inversely with $a_{PHY}(\lambda)$ peaking where $a_{PHY}(\lambda)$ was minimal (see Figure 3), this concurs with the findings of Bricaud and Morel (1983). The $\Delta\chi^*(\lambda)$ for *Anabaena* and *Oscillatoria* (Figure 9) and *Rhodomonas* (Figure 10) were influenced by $a_{PHY}(\lambda)$ to a lesser extent, while the $\Delta\chi^*(\lambda)$ for *Asterionella*, *Fragilaria*, *Euglena* and *Cryptomonas* (Figure 10) showed little spectral variation. Of the three Cyanophytes shown in Figure 9 *Microcystis* had a $\Delta\chi^*(\lambda)$ with an underlying trend which was inversely proportional to wavelength. A similar wavelength dependency for gas vacuolate Cyanophytes was measured by Davis-Colley *et al.* (1986) and Morel and Bricaud (1986).

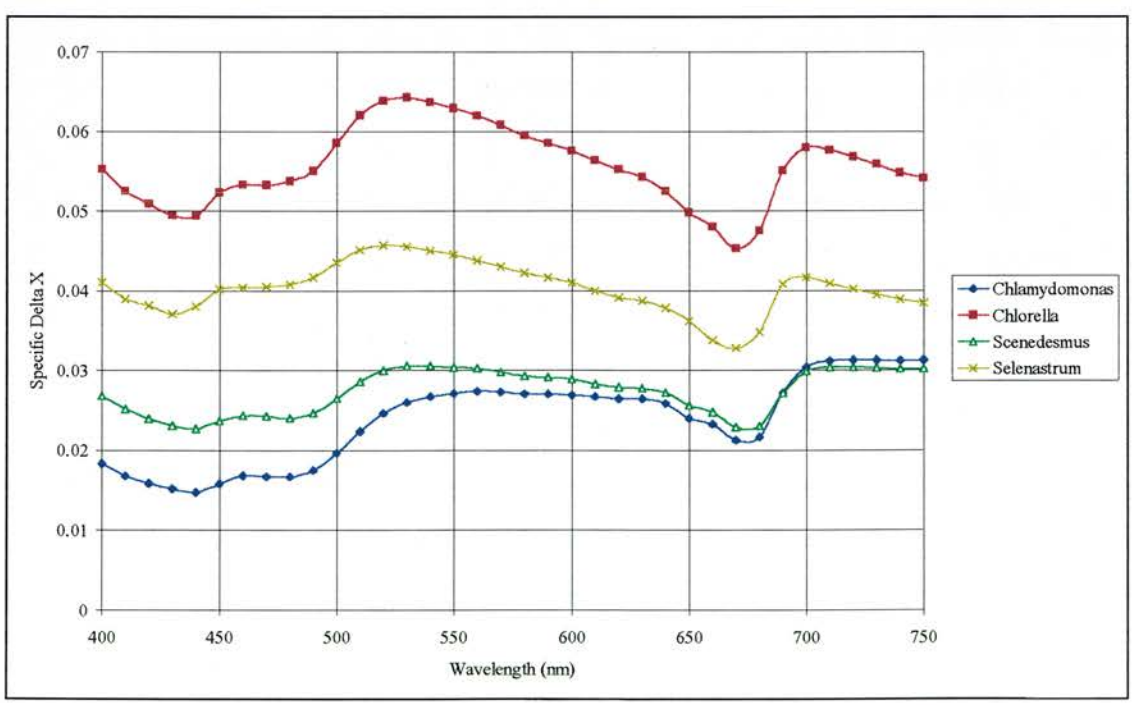

Figure 8. Pure culture Chlorophyte $\Delta\chi^*(\lambda)$.

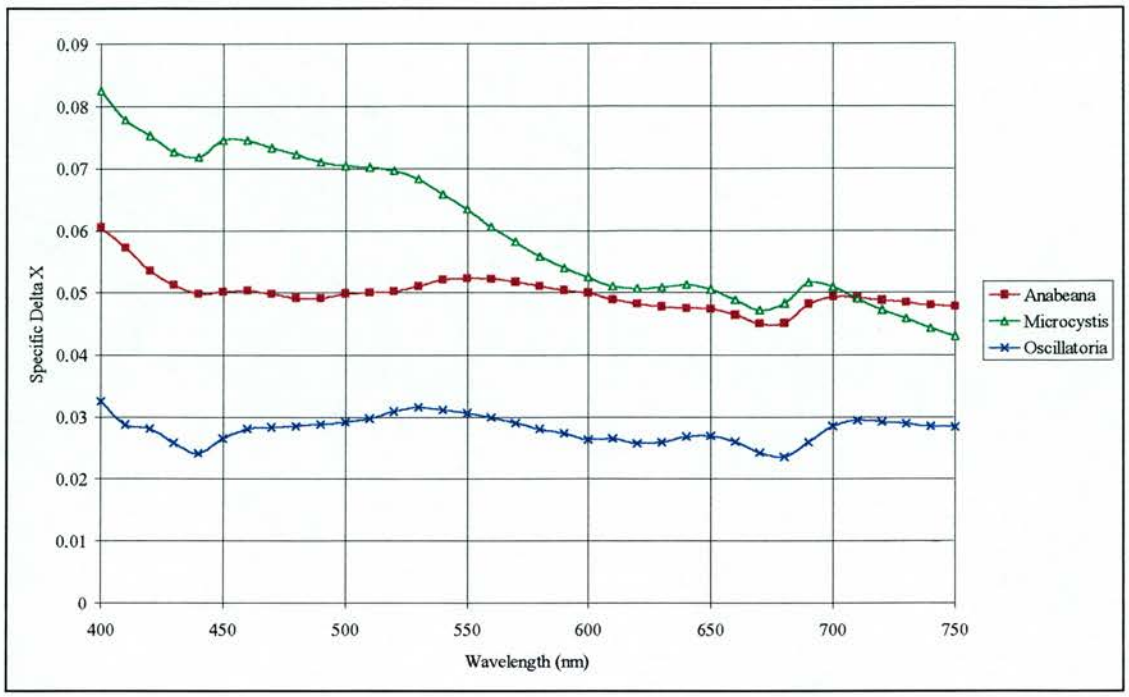

Figure 9. Pure culture Cyanophyte $\Delta\chi^*(\lambda)$.

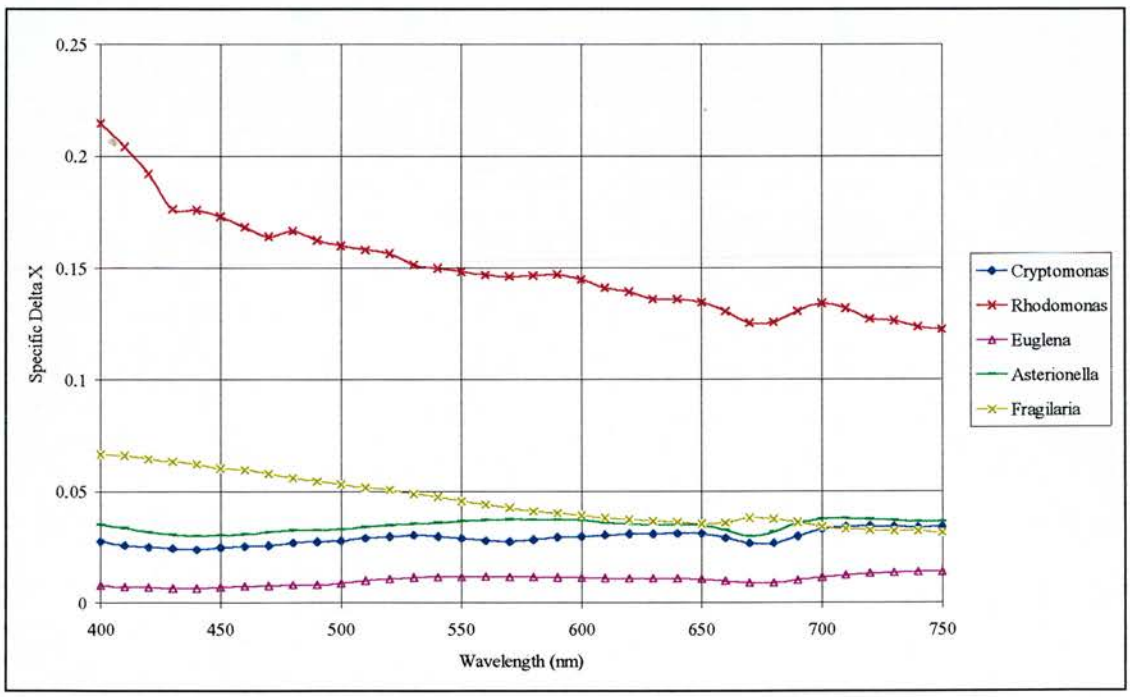

Figure 10. Pure culture Bacillariophyte (*Asterionella* and *Fragilaria*), Cryptophyte (*Cryptomonas* and *Rhodomonas*) and Euglenophyte $\Delta\chi^*(\lambda)$.

The inverse relationship of $\Delta\chi(\lambda)$ with $a_{PHY}(\lambda)$ for the larger celled species is predicted by the anomalous dispersion theory of Van de Hulst (Morel and Bricaud 1986). A spectrally trending $b_{PHY}(\lambda)$ will influence $R_{PHY}(\lambda)$ according to Equation 2.4.7. However this particular form of spectral variation

is of little significance to this study because as variation in $a_{PHY}(\lambda)$ is a function of pigment content, then variation in $b_{PHY}(\lambda)$ will also be a function of pigment content. The wavelength dependency of *Microcystis* arises because it is optically small (most probably it is the intracellular gas vacuoles that are responsible for most of this scattering). This form of scattering trend is important to this study because it is an influence upon $R_{PHY}(\lambda)$ caused by a phenomena which has no relation to pigment content hence no relationship to class, consequently it has the potential to reduce the effectiveness of class identification routines based on the recognition of class marker pigment features.

4.2.7 Phytoplankton volume scattering functions

Analytical interpretation of $R_{PHY}(\lambda)$ requires knowledge of all the inherent optical properties of phytoplankton, including the volume scattering function $\beta(\theta)$ which describes the angular distribution of scattered photons (Dekker 1993). The importance of $\beta(\theta)$ is that phytoplankton with relatively strong $b_{bPHY}(\lambda)$ will reflect proportionally more photons than those with relatively weak $b_{bPHY}(\lambda)$ which has implications for $R_{PHY}(\lambda)$ according to Equation 2.2.3.6.

$\beta(\theta)$ is determined by cell size, cell shape and the bulk refractive index. Phytoplankton cells are optically large producing a $\beta(\theta)$ which is strong in the forward direction and weak in the backward direction (Morel and Bricaud 1986). This pattern is enhanced as cell size increases which simultaneously increases forward scattering while narrowing the region of peak scatter, however increasing n or reducing cell size will increase $b_{bPHY}(\lambda)$. By virtue of random orientation the influence of cell shape on $\beta(\theta)$ remains minor for most of the common shapes exhibited by algae (Bricaud *et al.* 1988). Furthermore phytoplankton cells are not homogenous possessing complex internal cell structures which can complicate the scattering process.

Measurements of the angular scattering behaviour of pure cultures of phytoplankton were conducted at the Department of Physics and Astronomy in conjunction with the Institute for Environmental Studies, both at the Vrije Universiteit Amsterdam and was financed by The Netherlands Remote Sensing Programme (Dekker *et al.* 1997). There were several members in the team involved in the measurements of the angular scattering of phytoplankton (Volten *et al.* 1997). The experimental set-up shown in Figure 11 has previously been described in Kuik *et al.* (1991) for the measurement of atmospheric aerosols. The modifications the set-up required for the measurement of angular scattering from phytoplankton cells suspended in water are described in Volten *et al.* (1997).

A Helium-Neon laser beam at 633 nm was directed at a sample of pure culture phytoplankton suspended in high purity HPLC quality water and the light scattered from the sample was detected by a photomultiplier mounted on a goniometer ring. To ensure the absence of multiple scattering the maximum concentration within the single scattering range was ascertained by checking for deviation from linearity in scattering measured at a constant angle as cell concentration was increased. The

concentration immediately prior to the deviation from linearity was used. The sample cuvette was itself suspended in a basin filled with glycerine to minimise the changes in refractive indices occurring in proximity to the sample cuvette. The angular scattering distribution (ASD) was recorded at 5° intervals between 15° to 165°. The ASD is the scattered flux at each angle, to obtain the $\beta(\theta)$ it is necessary to divide the ASD by total scattering.

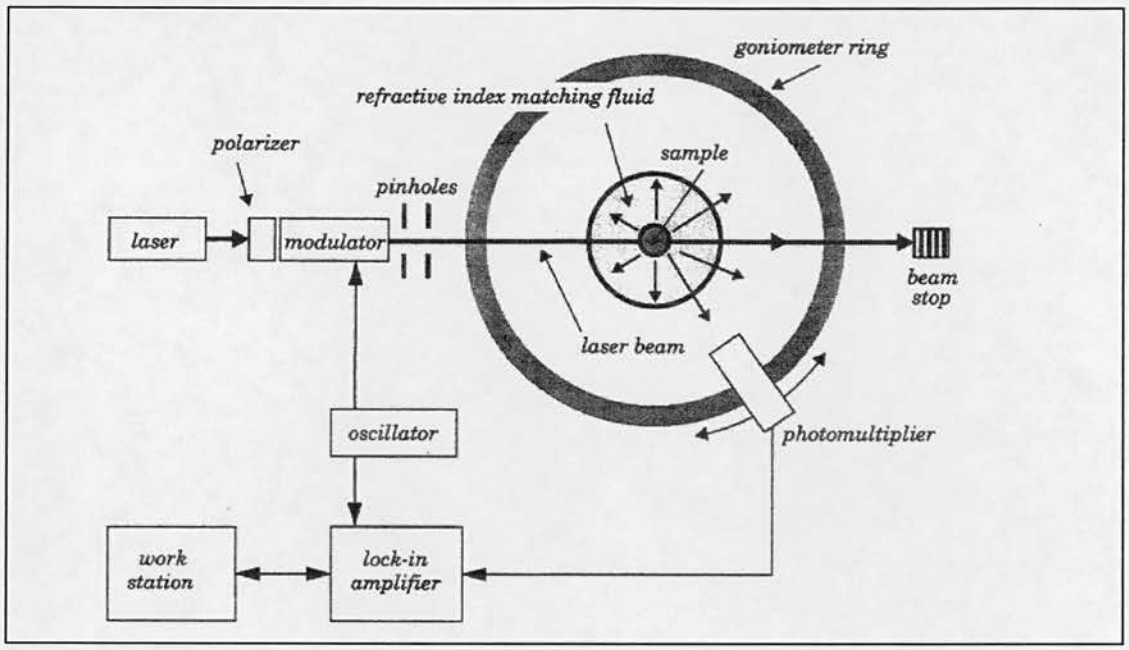

Figure 11. Schematic view of the experimental set-up for measuring the ASD showing laser, sample and photomultiplier. (Taken from Scheurs 1996).

The ASD as measured with the above set-up required correction to remove the four forms of experimental error (Scheurs 1996). A correction for the background scattering by the suspending water was required for each scattering function. A sine correction was required to account for the changes in the volume of observed sample caused by the changing viewing angle. A reflection correction was required to remove that portion of light not originally scattered towards the photomultiplier but eventually reaching the photomultiplier following various scattering and reflection episodes within the sample chamber. Finally correction for signal fluctuations was required.

Having corrected the ASD for these experimental errors it was necessary to fit the measured scattering functions to Mie functions in order to extrapolate to the narrow forward and backward angles necessary for the complete 180° scattering to be ascertained. Mie fitting functions are based upon complex (real and imaginary) refractive indices, cell size and cell size distributions. These input parameters were estimated or measured for each sample and then optimised to calculate the Mie scattering function which best fitted the measured data set for each sample. Goodness of fit of the calculated Mie function and the measured ASD was determined using least squares analysis.

The scattering probability curve for each angle describes the likelihood of a photon being scattered at a particular angle and can be calculated from the ASD for each alga. The importance of the shape of the scattering probability curve is its effect on R_{PHY}. R_{PHY} is increased for those cells with high b_{bPHY} and reduced for those cells with a relatively weak b_{bPHY}. *Aulacoseira granulata* (previously known as *Melosira granulata*) had the most dominant b_{fPHY} with 99% of scattered photons being scattered at angles between 0 and 5°. *Anabaena, Asterionella* and *Volvox* were also very dominant b_{fPHY} with 99% of scattering events being at angles between 0 and 20°, *Selenastrum* had a $\beta(\theta)$ which is slightly less peaked forward direction. *Oscillatoria* and *Microcystis* are relatively weak forward scatters with > 50% of scattering events being at angles between 0 and 20° consequently there is an increased likelihood that a scattered photon will be in the realm of b_b. A representative sample of the scattering probability curves for the freshwater phytoplankton used in this particular study are shown in Figure 12 along with the scattering probability for the $\beta(\theta)$ measured by Petzold (1972) in San Diego harbour. The San Diego harbour $\beta(\theta)$ has frequently been used by modellers of the underwater light climate including the influential work of Kirk (1981a, 1981b). The water sample is reported to be Case II in terms of its optical quality. The differences between the $\beta(\theta)$ of San Diego and *Microcystis* are obvious, imposing considerable implications for future modelling of the underwater light climate.

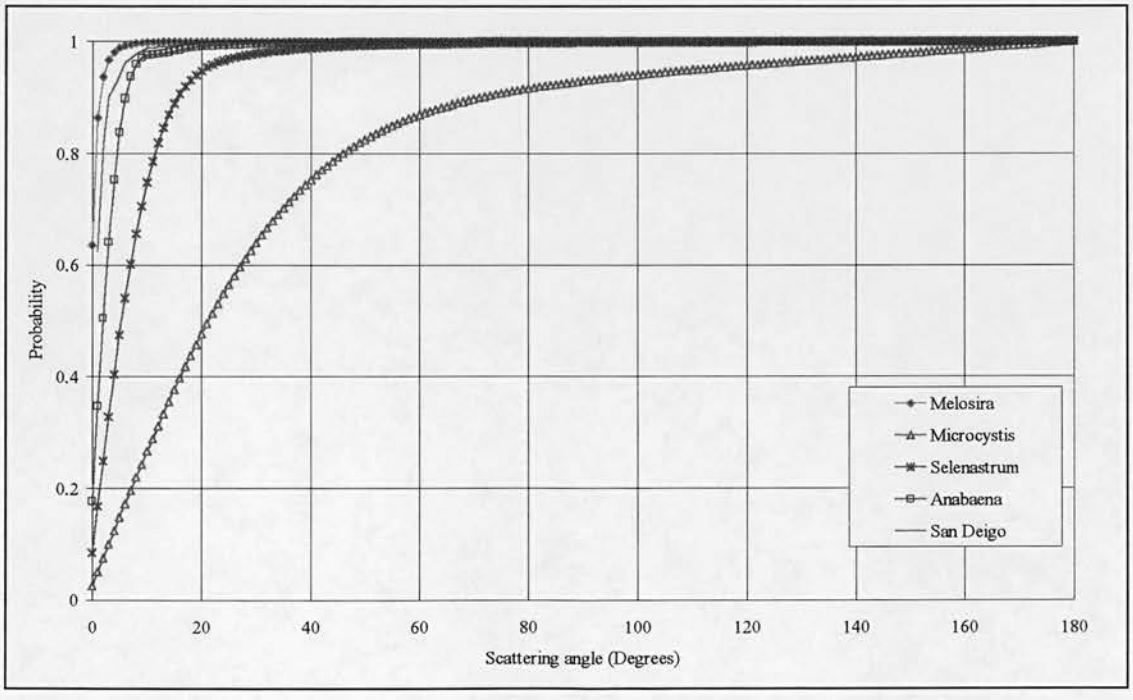

Figure 12. The scattering probability curves for pure culture phytoplankton. For comparison the scattering probability for San Diego harbour measured by Petzold (1974) (as reported in Kirk 1994) is also shown. The scattering angle is measured from the trajectory of the light beam had it passed unaltered through the scattering particle.

Analytical understanding of the variations in the ASD requires knowledge of the combined effect of cell morphology and the refractive index upon scattering. Large cell size tends to produce a relatively high forward scattering while backscattering increases with the refractive index (Morel and Bricaud 1986). Measurements of the refractive index of phytoplankton cells is difficult because various cellular components have to be considered (Aas 1981). Morel and Bricaud (1986) estimated refractive index from measurements of cell numbers, mean cell volume and absorption efficiencies and suggested that refractive index values fall within the following ranges: Bacillariophyte 1.009 to 1.024; Chlorophyte 1.023 to 1.039; and Cyanophyte 1.022 to 1.046.

Microcystis and *Selenastrum* are small single cellular species, consequently they scatter proportionally more in backward directions. The larger refractive index for the Cyanophyte further enhances this backward scattering. *Anabaena* and *Melosira* are larger cells and both form filamentous colonies, consequently they have proportionally more forward scattering, and for *Melosira* the dominance of forward scattering is further enhanced by a relatively low refractive index.

4.2.8 Summary of the optical properties of phytoplankton

This section has introduced the pure and mixed culture absorption spectra which are used in the following Chapters for the development of class identification routines. The inverse relationship between $a_{PHY}(\lambda)$ and $R_{PHY}(\lambda)$, and the importance of pigment composition to $a_{PHY}(\lambda)$ have been stressed, the corollary being that $R_{PHY}(\lambda)$ contains information regarding pigment composition and that pigment composition has a high degree of within-class consistency thereby lending itself to class identification routines. The general spectral conformity of all the phytoplankton absorption spectra has also been mentioned. This has implications for the identification of subtle class marker pigment absorption features not only in $a_{PHY}(\lambda)$ but also in $R_{PHY}(\lambda)$. The influence of $b_{PHY}(\lambda)$ on $R_{PHY}(\lambda)$ has been discussed in terms of the shape of the previously unmeasured $\beta(\theta)$ for pure cultures. Strong $b_{bPHY}(\lambda)$ leads to strong $R_{PHY}(\lambda)$ and is caused by cell morphology which has no within-class consistency so is not likely to be a source of information for class recognition routines.

4.3 Phytoplankton reflectance spectra

To develop algorithms and routines for the identification of phytoplankton classes from phytoplankton reflectance spectra ($R_{PHY}(\lambda)$) it is important to have some sample pure culture $R_{PHY}(\lambda)$ which are free from the influences of other optically active water quality parameters and from the adverse influences of the field environment. This section describes the laboratory measurement of such pure culture $R_{PHY}(\lambda)$ and discusses their salient features in terms of their absorption properties and pigment composition as identified in the preceding sections of this chapter. Within-class similarities and between-class differences in the $R_{PHY}(\lambda)$ are highlighted.

4.3.1 Performing reflectance measurements under controlled conditions

As was discussed in § 2.3.1 there are various drawbacks to the measurement of reflectance in the field which principally stem from variability of both the incident light field and of the components of the target water body. The effect of these drawbacks may be mitigated by performing the reflectance measurements in a controlled laboratory environment. This sub-section briefly reviews published work on the use of experimental laboratory tanks to make controlled reflectance measurements of aquatic targets, concentrating upon the particulars of the various tank designs.

The experimental tank dimensions are important for accurate simulation of the underwater light climate and subsequently for the accurate measurement of spectral signatures. Cylindrical tanks have been used to contain the suspended target in laboratory-based work conducted by Whitlock *et al.* (1982), Gons *et al.* (1992), Quibell (1992) and Han *et al.* (1994). Cubic tanks have also been used (Novo *et al.* 1989, 1991, Krijgsman 1994). Tanks must be sufficiently deep enough to be considered optically deep for a range of target suspensions. The upwelling radiance from optically deep waters is not influenced by bottom reflectance, conversely the upwelling radiance from optically shallow water bodies does contain a significant amount of radiance originating from bottom reflectance. Gordon and McCluney (1975) showed that most light reflected from within a water body originates from a depth corresponding to 1/Kd (§ 2.2) and thus is related to the optical properties of the water column.

Any experimental tank should have a large enough surface area to accommodate the area-of-view (AOV) of the spectroradiometer's viewing optics. AOV is dependent upon the field-of-view (FOV) and viewing distance. A combination of small surface area and wide FOV, or narrow FOV with large viewing distance will cause the tank walls to interfere with the signal. Typically narrow FOV ($< 6°$) have been used in combination with a small viewing distance (Whitlock *et al.* 1982, Quibell 1991).

Tank side-wall and base colour must also be considered. White side walls reflect photons diffusely thereby disturbing the angular distribution of the underwater light field which is responsible for the generation of the spectral signature. Most workers have opted for white walls and black bases to maximise the reflectance signal through multiple scattering from the side walls but minimising the potential for bottom reflectance (Novo *et al.* 1991, Quibell 1991, 1992). Krijgsman (1994) used mirrored side walls to fully develop the idea of an infinite horizontallity and reduce the alteration to the underwater radiance field imposed by reflectance properties of white side walls. Others, seeking to minimise the multiple scattering within their laboratory tanks, have painted the tank walls black which removes side-scattered photons from the radiance field thereby reducing the flux of the upwelling radiance (Mantovani and Cabral 1992, Han *et al.* 1994). To achieve a similar affect Novo *et al.* (1989) covered their tank walls in black painted sand.

Quibell (1991, 1992) and Han *et al.* (1994) placed their experimental tanks outside and relied upon a solar source; serendipitously both were located in sunny climates. Indoors, a suitable illumination source is required as a substitute for the solar source of the field environment. 1000 W tungsten-halogen lamps have been used to keep the signal-to-noise ratio high (Mantovani and Cabral 1992, Novo *et al.* 1989, Novo *et al.* 1991). The output spectra from these lamps is limited in the UV, rises through the visible and peaks in the near-IR and IR and as such is most unlike the solar spectrum. This has consequences for the spectral quality of the measured upwelling radiance, but the calculation of reflectance as a ratio of the upwelling to downwelling radiance flux removes the spectral influence of the downwelling irradiance. Typically direct illumination sources were favoured over diffuse sources (Novo *et al.* 1989).

Various spectroradiometers have been employed by workers measuring reflectance in controlled environments. Those opting for outside tanks have used CCD array type spectroradiometers like the Licor (Quibell 1992) and the Spectron SE-590 (Han *et al.* 1994). Those operating indoors are not restricted to such spectroradiometers as a constant illumination source facilitates the use of the typically more precise scanning spectroradiometers like the Macam SR-9010 (Gons *et al.* 1992).

4.3.1.1. Experimental tank design

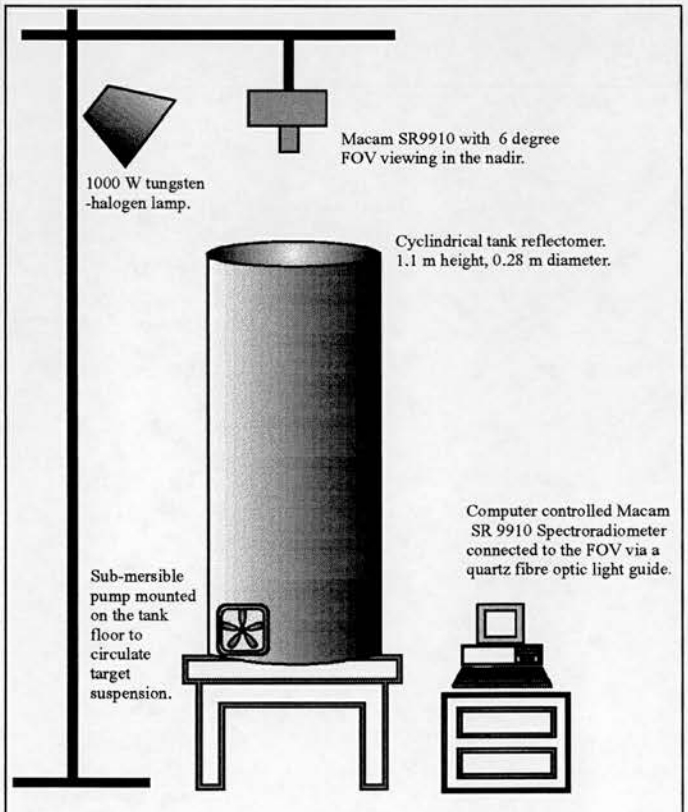

Figure 13. Schematic of the experimental tank design.

The experimental tank used in this study was a 1.1 m long vertically mounted cylinder with a 0.28 m internal diameter, giving a volume of 67.7 L (Figure 13). The side-walls were painted white and base black to maximise the return signal while creating an optically deep column. A metal frame was constructed above the tank on which the spectroradiometer optic and a 1000 W tungsten-halogen light source were mounted. The lamp was orientated to deliver light at 45° to the water surface. The spectroradiometer optic had a 6° field-of-view (FOV), so when viewing in the nadir and suspended at 0.35 m above the water's surface it had an AOV at the top of the tank of 0.001 m^2. In this position the tank walls would not interfere with the upwelling radiance measurements. To

reduce the generation of a diffuse irradiance flux within the laboratory the frame was used to support a black-out cloth and the only light source in the laboratory was the tungsten-halogen lamp. A black painted submersible pump was placed on the floor of the tank to create an internal circulation to maintain a homogeneous target suspension.

Radiance measurements were performed with a Macam SR 9910 scanning spectroradiometer (Macam, Livingston, Scotland). The SR 9910 is a precise and accurate scanning spectroradiometer. Fitted with a double concave holographic grating monochromator it achieves a spectral resolution of ± 0.5 nm at wavelength steps of ± 0.25 nm over a spectral range of 240 to 800 nm. For this study all spectral measurements were performed in triplicate at 1 nm intervals from 380 to 800 nm and the average value used in further analysis. The multi-alkali photomultiplier tube is fitted with a stabilised power supply. The signal is amplified by a six decade transimpedance amplifier. The gain across the amplifier is automatically checked prior to each wavelength increment and measurement. Each recorded measurement is the average of a user-defined number of analog-to-digital (ATD) conversions ranging from 250 to 1000. For the experimental tank work this was set to 650 ATD. A dark current measurement is automatically made prior to each scan to correct for residual current within the transimpedance amplifier.

The SR 9910 was controlled through a portable computer via an RS232C interface enabling the user to control many aspects of the instrument set-up before each scan. Spectra were automatically calibrated against calibration files and then stored to disk. Calibration of the optics was performed by the supplier against a tungsten halogen lamp calibrated to National Physics Laboratory (UK) standard. Re-calibrations of the spectrophotometer were made quarterly.

For measurements of pure phytoplankton cultures the tank was filled with sand-filtered water into which the phytoplankton were systematically added to achieve nine stepped increases in concentration. The different concentrations were selected to simulate a range of water types from oligotrophic to hyper-eutrophic. Two species from three different classes were used: *Microcystis* and *Anabaena*; *Selenastrum* and *Chlorella*; and *Asterionella* and *Fragilaria* (from the classes Cyanophyceae, Chlorophyceae and Bacillariophyceae respectively). To assist in the interpretation of the reflectance measurements they were supplemented with CHL*a* analysis, measurements of $a_{PHY}(\lambda)$ and $\Delta\chi(\lambda)$, and cell counts and cell size measurements.

The radiometric measurements required to calculate reflectance from the experimental tank are described below. The preferred form of reflectance was below-surface irradiance reflectance $R(\lambda)(-0.01m)$ because this is least affected by specularly reflected radiance (§ 2.2.3). The $Lu_a(\lambda)$ flux was measured with the 6° FOV viewing in the nadir 0.35 m above the centre of the tank. The angle of view was checked with a fish-eye spirit level. $Lu_a(\lambda)$ required correction for the fraction of

downwelling radiance specularly reflected from the water surface. Due to the 45° illumination angle of the tungsten-halogen lamp it may be assumed that specularly reflected photons from the surface of the water were reflected at approximately 45° and consequently would not have reached the optic. However within the laboratory environment there was a small component of diffuse irradiance which may have been specularly reflected into the FOV and recorded as part of upwelling signal. The correction of this diffuse component is described in § 2.2.3 after Dekker (1993). The method requires the ratio of diffuse to direct downwelling irradiance $F(\lambda)$. $F(\lambda)$ was calculated by dividing downwelling irradiance (as measured from a reference panel held perpendicular to the nadir immediately above the water's surface) by the radiance measured at +180° to the lamp's azimuth at the zenith angle of the lamp. For the experimental tank $F(\lambda)$ was found to be very low (Figure 14). $Lu_a(\lambda)$ was corrected accordingly and was then converted into $Eu_w(\lambda)$. If water had Lambertian properties then $Eu(\lambda) = Lu(\lambda)\pi$. However the underwater light climate is not Lambertian and the conversion factor Q is required to convert from radiance to irradiance (§ 2.2.4). Given that Q has an influence upon the magnitude of the calculated $Eu_w(\lambda)$ but not on its spectral quality (Dekker 1993), the debate concerning the actual value of Q was not considered applicable to this study so Q was set to 5.

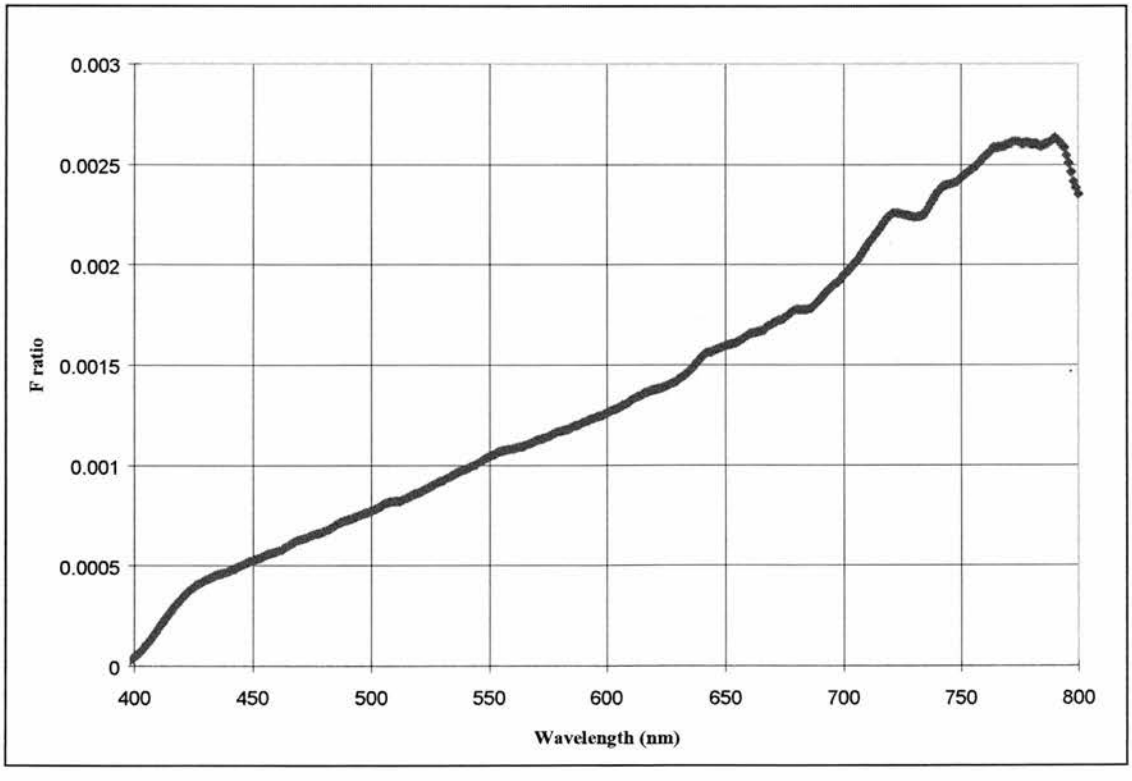

Figure 14. The ratio of direct to diffuse irradiance ($F(\lambda)$) for the experimental tank environment.

A measurement of $Ed_w(\lambda)$ was then required to calculate $R(\lambda)(-0.01\ m)$. $Ed_w(\lambda)$ was measured with a cosine correcting head placed 0.01 m below the surface of the water. Although this is a direct measure of $Ed_w(\lambda)$ the physical size of the cosine correcting optic prevented a small fraction of the upwelling

irradiance flux from reaching the water's surface, a percentage of this flux would have been reflected from the air-surface interface and added to downwelling flux. Consequently this measurement could not be considered entirely accurate. $R(\lambda)(-0.01m)$ was calculated according to Equation 2.2.3.5.

Given that the experimental tank is an artificial environment in which to place target samples for the purpose of measuring their reflectance properties it is important to consider how the tank design affects those reflectance properties.

4.3.1.2 The experimental tank illumination source

The output from the tungsten-halogen lamp increases proportionally with wavelength being relatively low in the blue region of the visible spectrum and steadily increasing to the near-IR (Figure 15). The low output flux in the shorter wavelengths has implications because of the reduced signal to noise ratio. In addition the PAR integrated surface incident irradiance flux measured at top of the tank wavelengths was 37.87 Wm^{-2} which is low compared to the typical mid-day summer incident solar flux measured on both cloudy and sunny days at the latitude of the English Lake District which was 135 Wm^{-2} and 290 Wm^{-2} respectively.

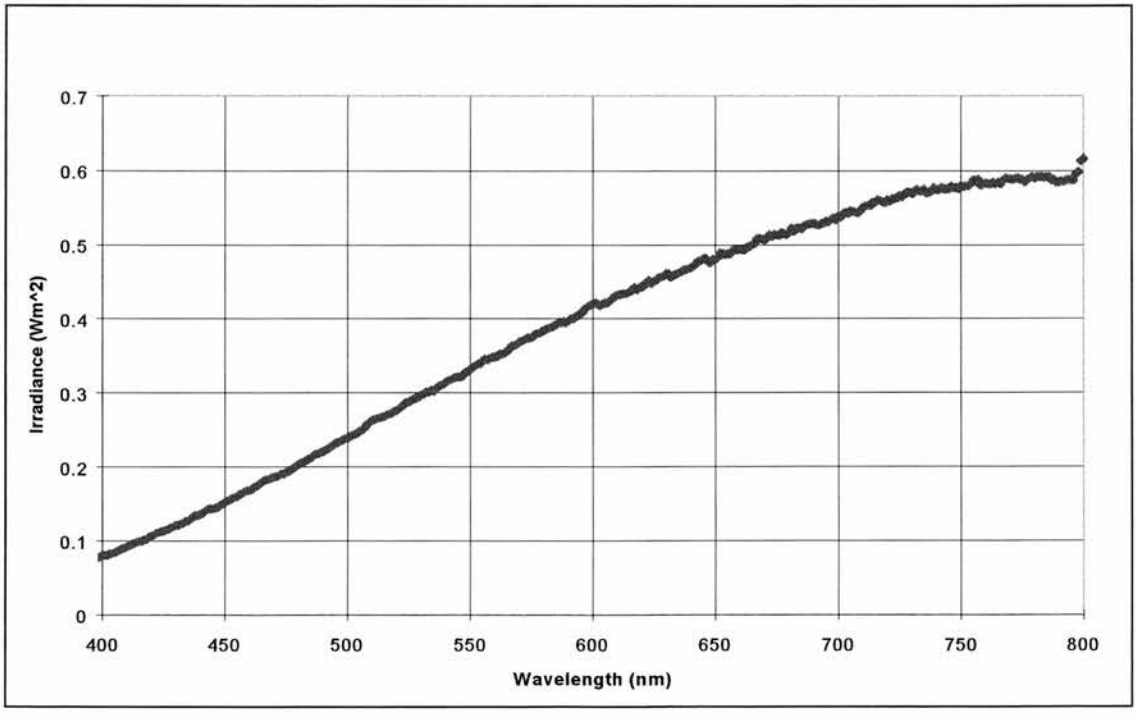

Figure 15. Example spectrum from the 1000 W tungsten-halogen lamp.

The consistency of the tungsten-halogen output is also important because the time required to make a single scan from 400 to 800 nm at 1 nm intervals using the SR 9910 was *ca.* 7 min. Figure 16 shows the variation in the lamp's output during a single scan as measured by a cosine correcting PAR sensor. Summary statistics for PAR meter readings made over a period of 147 min are: a mean flux

density of 170.0 Wm^{-2} ; a standard deviation of 3.5; and a variance of 12.2. Both these measurements suggest that the output was relatively consistent over time.

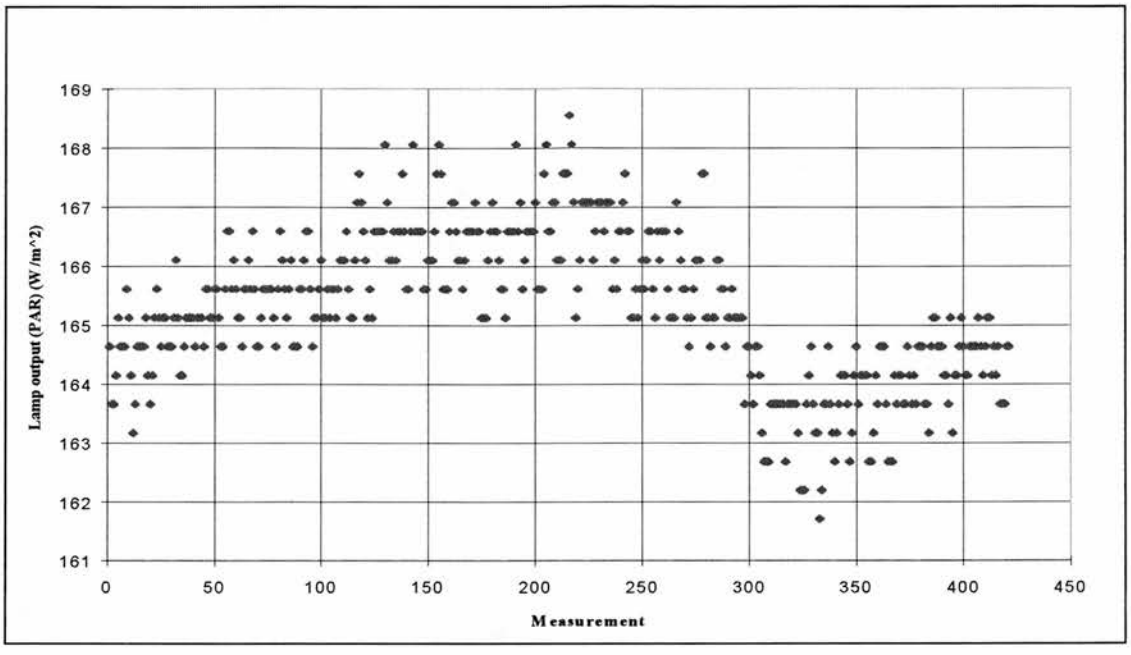

Figure 16. An example the variation in irradiance as measured by the PAR over a period of 7 minutes. Each measurement equates to approximately 1 second. Note the exaggeration of the *y*-axis.

The lamp was shone at 45° to the water's surface. The angle caused the lamp to cast a spatially heterogeneous irradiance flux across the water surface (Figure 17). This spatial heterogeneity is artificial and would not be present in the field. It could be responsible for variations in the magnitude of the $Lu_w(\lambda)$ generated from different areas of the tank. The larger the AOV the greater the degree of spatial variability in $Lu_w(\lambda)$ and the less relevant the measurement of $Ed_w(\lambda)$ in the centre of the tank. The consequences for this inhomogeneity will therefore be inaccuracy in the magnitude of the two spectra required for the calculation of the spectral signature. Figure 17 shows that the degree of spatial variation across the 0.001m^2 AOV of the 6° FOV optic was minimal.

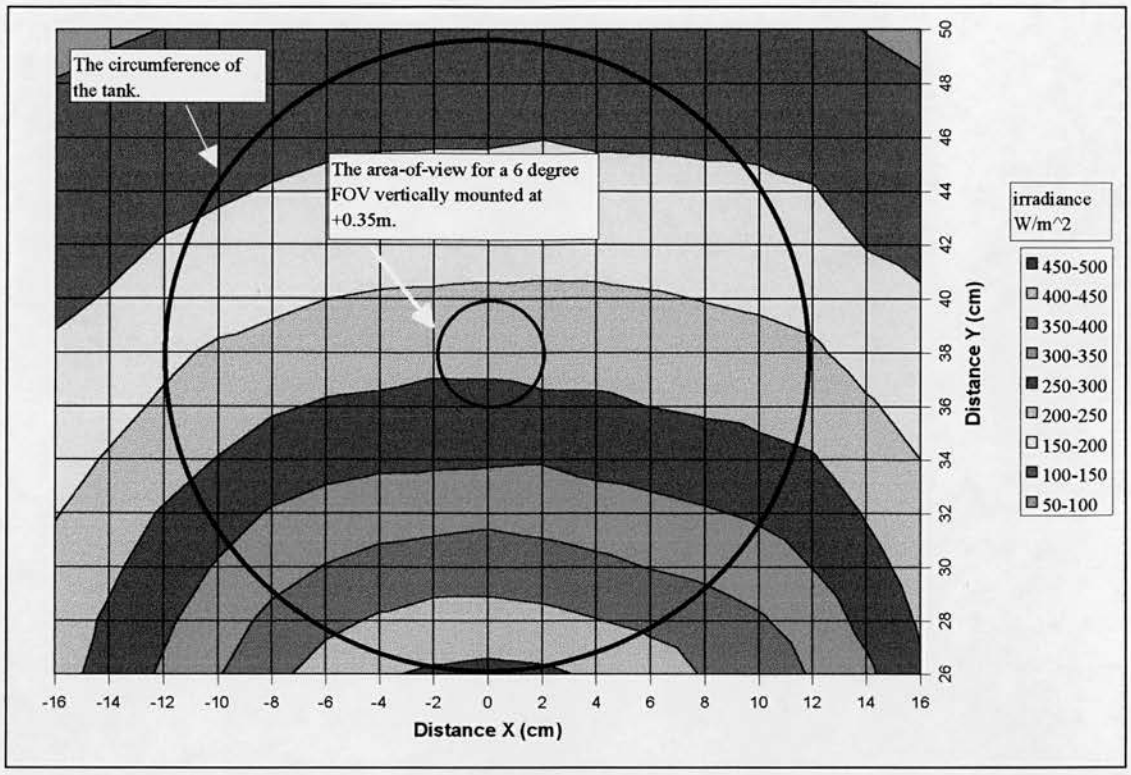

Figure 17. The downwelling PAR irradiance flux (Wm^{-2}) distribution across the surface of the tank. The measurements were made by placing a flat board marked with a 1 by 1 cm grid across the tank top, the PAR sensor was placed at each of the intersecting grid-lines and a measurement made. The circumference of the tank has been superimposed upon the graph surface. Distance X represents perpendicular distance from a line drawn parallel to the lamp's azimuth angle. Distance Y represents the horizontal distance from the lamp.

4.3.1.3 Experimental tank shadow

A further aspect of the tank's optical environment which required investigation originated from the 45° angle of illumination from the tungsten-halogen lamp. This caused the side-walls of the tank to create a shadow within the tank. The vertical extent of this shadow is present in the downwelling irradiance profile measured using a PAR cosine correcting sensor at the mid-point of the tank at 19 incremental depths when the tank was filled with sand-filtered water (Figure 18). The irradiance profile shows a reduction in irradiance with depth not consistent with the exponential decrease that would be expected in natural water bodies (Kirk 1994). Deviations from a straight line in plots of $\ln(Ed_w)$ against depth (z) (*i.e.* the downwelling attenuation coefficient Kd) indicate that the medium was either not homogeneously mixed or that the radiance distribution of the underwater light climate changed with depth. The values of Kd (PAR) for the above and below shadow region further supports this anomaly with the change from a direct to a more diffuse light field occurring at *ca.* -0.3 m (Table 4).

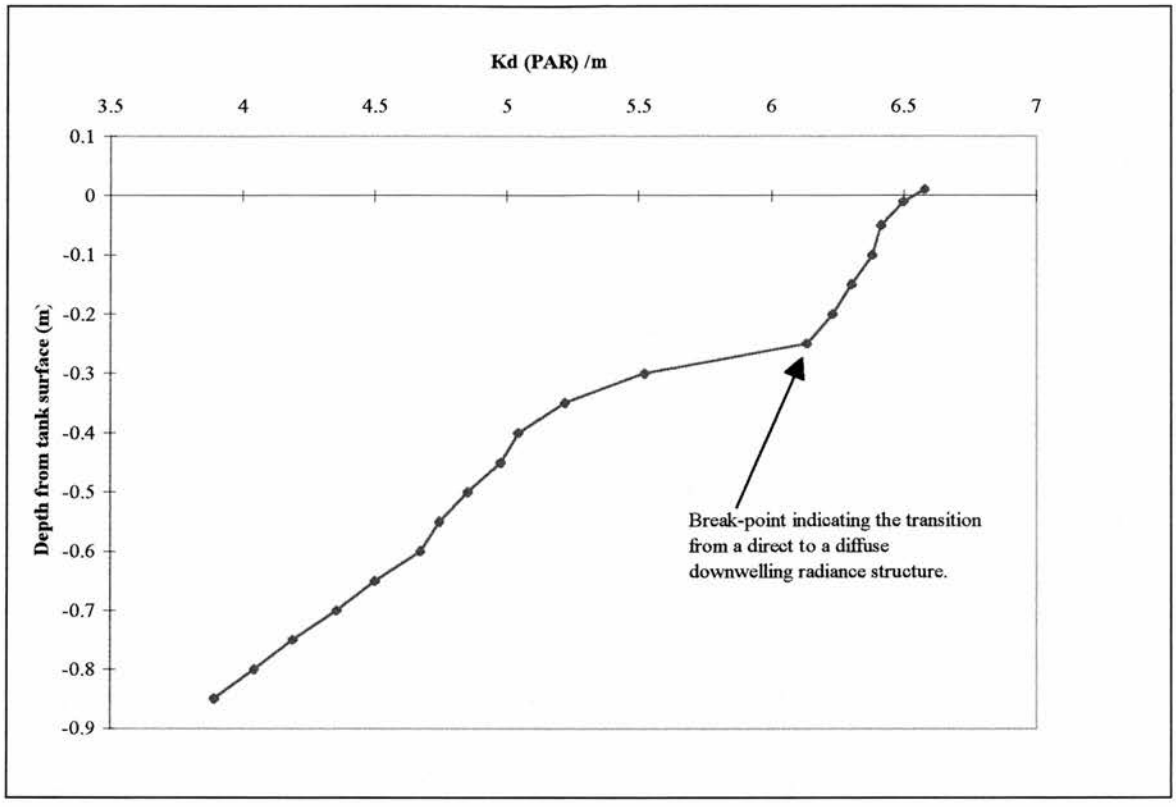

Figure 18. The Kd(PAR) profile for the experimental tank. The combination of the angle of illumination and the tank side-walls causes a shadow to be cast changing the underwater radiance distribution from direct to more diffuse thereby altering the value of the Kd(PAR).

Table 4. The Kd(PAR) for the above and below shadow regions of the tank. The shadow region was extrapolated from the breakpoint in the plot of $\ln(Ed_w)$ against depth (Figure 18).

Tank depth	*Attenuation coefficient*	r^2
Kd(PAR) above shadow (-0.01 to -0.25 m)	1.4	0.986
Kd(PAR) below shadow (-0.35 to -0.85 m)	2.6	0.990

The effect of the tank shadow on the spectral signature is likely to be limited. There will be an effect upon the generation of the upwelling radiance flux as the underwater light climate changes from a predominantly directional to a predominantly diffuse flux, however this flux change would be difficult to quantify as it would vary with the optical properties of the target suspensoids. The benefits of the 45° illumination angle outweigh the uncertainties of the errors arising from tank shadow. Quibell (1992) also noted the presence of tank shadow caused by a 45° angle of illumination and reached the conclusion that the reflectance spectra obtained from the system were only suitable for qualitative analysis. No further investigations into this effect were conducted.

4.3.1.4 Modelling the influence of experimental tank depth upon the reflected flux

The tank bottom was painted black to give an optically deep effect, however the physical depth of 1.1 m was likely to be restrictive when the optical depth of the target suspension was low. The effect of the restricted tank depth was assessed with the use of a Monte-Carlo radiative transfer model

developed by Harwar (1995) and modified for this purpose. The model tracked the progress of 25,000 photons as they entered a hypothetical water-body at 45°. The depth of the hypothetical water body was set to 10.0 m. Any photon reaching this depth was 'absorbed'. The inherent optical properties for those phytoplankton species used in the experimental tank were used as input into the model. The model identified the wavelength with the lowest total attenuation wavelength $c(\lambda_{min})$ as this was the wavelength with the shallowest optical depth. Photons at this wavelengths should penetrate the deepest into a water body hence are most likely to reach the tank floor. The model followed the progress of photons at this wavelength and those which reached a depth equivalent to the bottom of the experimental tank (1.1 m) were tagged. Of these photons, the number which were subsequently reflected were tagged again, thereby allowing the percentage of photons being reflected from depths below 1.1 m to be determined. Table 5 shows the percentage of photons that would have attained depths below 1.1 m yet would still have been reflected had the tank been 10 m not 1.1 m deep.

Table 5. The summary output from the Monte-Carlo modelling of the experimental tank. The model was designed to assess the effect of restricted tank depth on the reflected flux and is described in the text. The effect for the strongest and weakest concentration of each species used in the experimental tank was modelled at the $c(\lambda_{min})$ wavelength.

Species and series identifier	*CHLa* (μgL^{-1})	$c(\lambda_{min})$ *(nm)*	$c(\lambda_{min})$ (m^{-1})	*% photons attaining a depth of > 1.1 m which were subsequently reflected back through the water's surface*
Anabaena 1	9.5	568	0.53	2.5
Anabaena 9	126.3	712	5.86	0.1
Asterionella 1	8.8	432	0.46	1.5
Asterionella 9	53.2	664	2.52	0.3
Chlorella 1	12.1	572	0.99	3.7
Chlorella 9	89.5	668	6.31	0.1
Fragilaria 1	3.9	492	0.24	1.3
Fragilaria 9	30.5	584	1.35	0.7
Microcystis 1	15.6	580	2.39	2.9
Microcystis 9	103.9	720	12.39	0.2
Selenastrum 1	30.7	576	1.09	4.6
Selenastrum 9	183.9	656	5.61	0.1

It can be seen from Table 5 that at the highest concentration of each species the reduction in reflectance due to the restricted tank was minimal. At the lower concentrations the tank depth had much more of an influence. Interestingly, the extent of this influence does not appear to be directly related to the magnitude of the beam attenuation coefficient with *Selenastrum* at 30.7 mg CHLa L^{-1} losing 4.6% of the reflected flux to the tank base while the *Fragilaria* at 3.9 mg CHLa L^{-1} only lost 1.3%. This contradictory result is due to the stronger scattering coefficient of the *Selenastrum* resulting in more scattering events and therefore more reflectance; although more photons will attain > 1.1 m depth in the *Fragilaria* sample, very few would subsequently be scattered back to the surface.

Modelling the effect of the restricted tank depth on the generation of reflectance was necessary because photons reaching the black tank base will be absorbed, hence lost from the underwater light

climate, and will therefore be unavailable for the generation of an upwelling flux. The modelling exercise has shown that the restricted depth does reduce reflectance and the magnitude of this reduction is related to the strength of the scattering coefficient. Given that the model requires inputs of inherent optical properties which could in most instances only be approximated the model output should be considered relatively rather than absolutely. However the modelling has shown that the overall reduction in reflectance caused by the physical depth of the tank is low, especially at the higher concentrations, and may be ignored for the rest of the study.

4.3.1.5 Phytoplankton homogeneity within the experimental tank

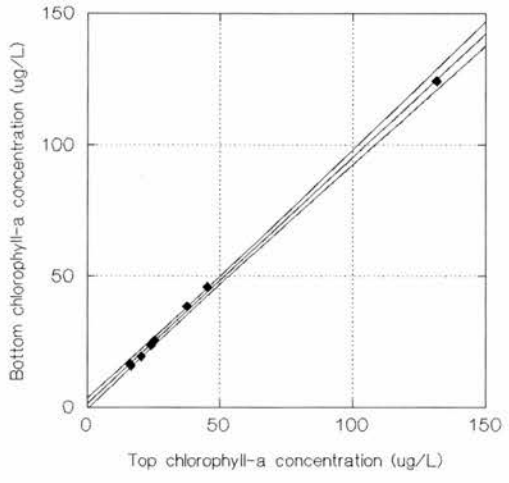

Figure 19. A comparison of the mean top and bottom chlorophyll-*a* concentrations in the experimental tank. The regression line for top against bottom chlorophyll-*a* concentration has the following characteristics, Bottom chlorophyll = 0.94 Top chlorophyll + 1.25 (r^2=0.999). Also shown are the 95% confidence limits for the regression line.

A black submersible pump was placed on the floor of the tank to generate an internal circulation to ensure that a homogeneous suspension was maintained. The ability of this pump to maintain a homogeneous suspension of phytoplankton was assessed by comparing the measured CHL*a* concentrations from samples taken contemporaneously from the top and bottom of the tank during a series of reflectance measurements. Three samples from each location were measured for CHL*a* concentration and the means have been plotted against each other in Figure 19. The slope parameter is close to unity suggesting that the suspension within the tank was homogenous. The deviation from a slope of unity indicated that there was no significant difference between the mean CHL*a* concentrations measured from the top and bottom of the tank while the pump was active (two-tailed t-test gave t = 1.4 with a t_{crit} = 2.4 for 8 observations).

4.3.2 Culture cell numbers and sizes

To complement the optical measurements, counts of the cell numbers for the cultures used in the experimental tank were made and average cell dimensions were also measured. The cell numbers for the pure culture prior to dilution in the experimental tank are shown in Table 6. Cell numbers for the various concentrations of the dilution series can be calculated using a simple mass balance based upon the measured CHL*a* concentration for each member of the series. The concentration of CHL*a* per cell shows that the two Bacillariophytes have considerably more CHL*a* per cell than do the Chlorophytes

and Cyanophytes. The average cell volume is also shown in Table 6 with the diameter of a equivalent spherical cell. The two Bacillariophytes are considerably larger than the cells from the other classes.

Table 6. Summary of cell counts and cell volume measurements. The intracellular pigment concentration and the diameter of the equivalent sphere are used in the calculation of the absorption by solution (§ 4.3.3.2).

Species	*Pure culture CHLa ($\mu g\ L^{-1}$)*	*No. of cells in pure culture (mL^{-1})*	*Intracellular CHLa (ng)*	*Assumed cell shape*	*Average cell volume (μm^3)*	*Diameter of the equivalent sphere (μm)*
Selenastrum	1947.3	2585948	0.8	Bi-cone	92.6 ± 6.6	11.2
Chlorella	819.4	5002140	0.2	Sphere	16.0 ± 1.9	5.7
Anabaena	1201.5	1359354	0.9	Cylinder	135.9 ± 8.6	12.8
Microcystis	1069.5	15200839	0.1	Sphere	19.0 ± 1.8	6.4
Asterionella	714.1	145745	4.9	Cylinder	398.5 ± 35.8	18.3
Fragilaria	320.6	96311	3.3	Cylinder	1092.6 ± 91.3	12.8

4.3.3 The inherent optical properties of the phytoplankton species used in the experimental tank

As discussed in § 4.2, and in a similar manner to the approach adopted by Malthus and Dekker (1990), measurement of some of the inherent optical properties of the species used in the experimental tank were required to facilitate a fuller understanding of the measured $R_{PHY}(\lambda)$.

4.3.3.1 The specific absorption spectra

As discussed in § 2.1.5.2 comparisons of $a_{PHY}(\lambda)$ at different biomass concentrations are best made by using the specific absorption spectra $a_{PHY}^*(\lambda)$. $a_{PHY}^*(\lambda)$ for the species used in the experimental tank and calculated according to Equation 2.1.5.2.2 are shown in Figure 20. It can be seen that although there are marked differences between the different species there are some broad similarities. The absorption peaks at 440 and 675 nm are consistently present in all spectra, as is the absorption minima at 550 nm, and the absence of absorption beyond 720 nm. The 440 and 675 nm absorption peaks are the Soret and far-red CHL*a* absorption peaks and are present in all the spectra because CHL*a* is a ubiquitous pigment, there are however slight variations in the location of the far-red CHL*a* absorption peaks for each species which occur at: 680 nm (*Selenastrum*); 674 nm (*Chlorella*); 680 nm (*Anabaena*); 676 nm (*Microcystis*); 676 nm (*Fragilaria*); and 680 nm (*Asterionella*). The spectral differences between $a_{PHY}^*(\lambda)$ for the different species arise from two sources: cellular structure and pigment content. The spectral implications of both are discussed below. The pigment compositions of each species have been identified from the literature and the $a_{PHY}^*(\lambda)$ are discussed in relation to these compositions.

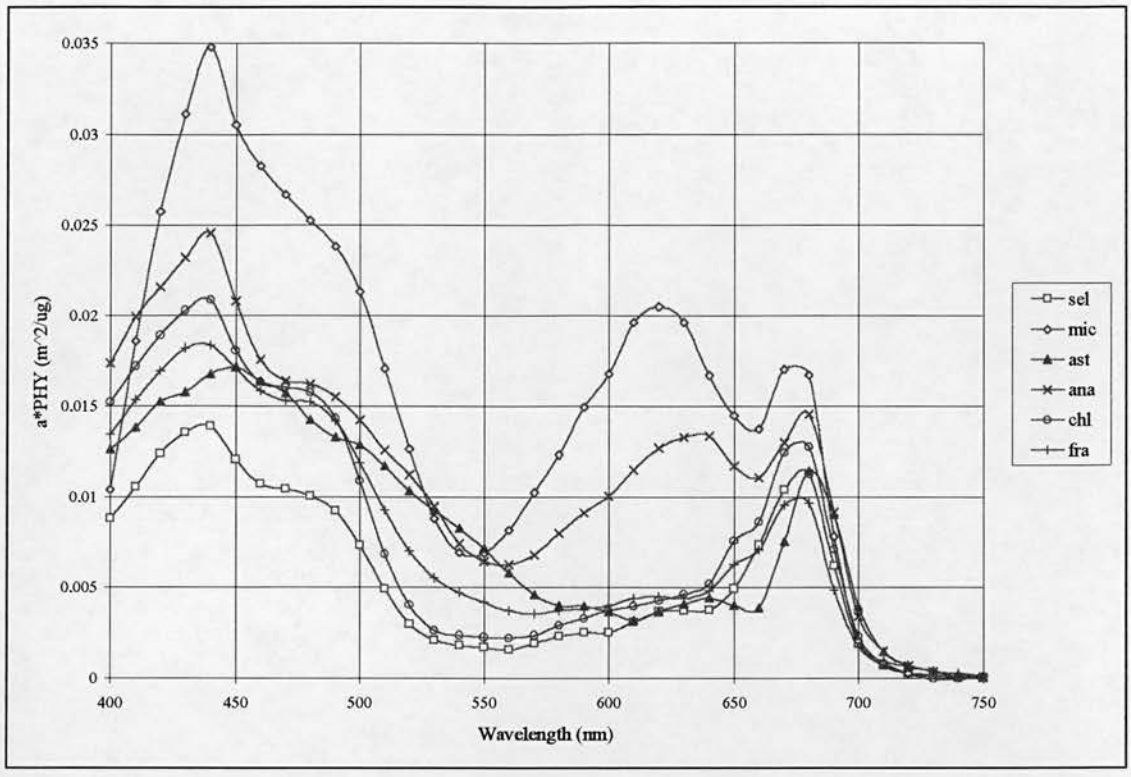

Figure 20. The $a_{PHY}^*(\lambda)$ for the pure culture phytoplankton measured in the experimental tank. In the legend sel is *Selenastrum*, mic is *Microcystis*, ast is *Asterionella*, ana is *Anabaena*, chl is *Chlorella* and fra is *Fragilaria*.

The two Chlorophytes principally contain chlorophylls, the presence of which can clearly be identified in the $a_{PHY}^*(\lambda)$ for *Chlorella* and *Selenastrum*. There is considerable between-species similarity in the measured $a_{PHY}^*(\lambda)$ with the following features being visible in both species: 440 nm (CHL*a*), 480 nm (CHL*b*), 625 nm (CHL*a*), 650 nm (CHL*b*)and 675 nm (CHL*a*). There is also an absorption feature at 580 nm which is most likely a biliprotein. In the shorter wavelengths *Chlorella* has more absorption per unit CHL*a* than does *Selenastrum*, but from 520 nm into the longer wavelengths the specific absorption coefficients differ little.

The two Cyanophytes principally contain CHL*a* and biliproteins. As with the two Chlorophytes these can be clearly identified in the $a_{PHY}^*(\lambda)$ for *Microcystis* and *Anabaena*. The following spectral features are common to both species: 440 nm (CHL*a*), 570 nm (PCE), 625 nm (PCC), 670 nm (APC) and 675 nm (CHL*a*), *Microcystis* also exhibits features at 495 nm (PCE) and 585 nm (PCC), *Anabaena* exhibits 420 nm (CHL*a*), 475 and 520 nm (non-specific carotenoids) and 640 nm (PCC). The biliprotein absorption features at *ca.* 625 nm is a very significantly feature, indeed for *Microcystis* the *ca.* 625 nm absorption exceeds the far-red CHL*a* absorption. The Cyanophytes exhibit the strongest $a_{PHY}^*(\lambda)$ with *Microcystis* absorbing twice as strongly at 440 nm than does *Selenastrum*.

The two Bacillariophytes principally contain chlorophylls and xanthinols. $a_{PHY}^*(\lambda)$ features common to both *Asterionella* and *Fragilaria* are 440 nm (CHL*a*), 620 nm (CHL*c*) and 675 nm (CHL*a*). Features restricted to *Asterionella* are 420 nm (CHL*a*), 450 nm (CHL*c*), 460 nm (FCX, DDX or CHL*c*), 500 nm (FCX), 540 nm (FCX), 590 nm (CHL*c*) and 640 nm (CHL*c*). The features restricted to *Fragilaria* are 480 nm (FCX) and 620 nm (CHL*c*).

4.3.3.2 Comparisons of specific absorption spectra in the absence of the package effect

In the absence of a package effect the variability in the $a_{PHY}^*(\lambda)$ can be taken to reflect variability in pigment composition. The package effect (Duysens 1956, Kirk 1981a, 1981b), described in § 2.1.5.2, arises because of the packaging of pigments into discrete clusters (on both a cellular and intra-cellular level). The package effect reduces the absorption efficiency of a suspension of phytoplankton cells when compared to the absorption by a hypothetical solution containing the same number of pigmented molecules in homogenous dispersions. The absorption efficiency is reduced because of the mutual shading of the pigmented molecules. The nature of the package effect is such that it is strongest when absorption is strongest (Duysens 1956). The degree to which the package effect flattens $a_{PHY}(\lambda)$ is dependent upon cell morphology.

The package effect can be eliminated from the *in vivo* $a_{PHY}^*(\lambda)$ by considering the absorption of an equivalent solution, $a_{SOL}^*(\lambda)$ (Bricaud and Morel 1983). Comparisons of the $a_{SOL}^*(\lambda)$ for the different species should highlight differences attributable solely to variation in pigment composition. Sathyendranath *et al.* (1987) provided a method for the calculation of the $a_{SOL}^*(\lambda)$ from $a_{PHY}^*(\lambda)$ which was applied to the $a_{PHY}(\lambda)$ presented above.

$a_{SOL}^*(\lambda)$ was calculated according to the successive approximation method of Sathyendrananth *et al.* (1987) which was adapted for this study to include the whole of the visible spectra. A Turbo Pascal v7.0 program (*Particle.pas*) was written for this purpose. Using the measured $a_{PHY}(\lambda)$, cell counts and cell size measurements (converted to the spherical equivalent) then the absorption efficiency $Q_a(\lambda)$ can be calculated according to:

$$Q_a(\lambda) = \frac{4a_{PHY}(\lambda)}{\pi d^2 N} \quad \text{.....4.3.3.2.1}$$

where d is the diameter of the equivalent sphere and N the number of cells all of which had been measured. $Q_a(\lambda)$ can also be calculated according to:

$$Q_a(\lambda) = 1 + \frac{2\exp(-\rho)}{\rho} + \frac{2(\exp(-\rho)-1)}{\rho^2} \quad \text{.....4.3.3.2.2}$$

where ρ is a cell size parameter. Because Equation 4.3.3.2.2 is a monotonic relationship ρ can be solved by successive approximation. Having obtained ρ the absorption attributable to the cellular material $a_{CM}(\lambda)$ can be calculated according to:

$$a_{CM}(\lambda) = \rho / d \quad \text{.....4.3.3.2.3}$$

Were the cellular material to be evenly dispersed in a suspension then it would be equivalent to $a_{SOL}(\lambda)$. Absorption by such a homogenous suspension can be calculated using:

$$a_{SOL}(\lambda) = (\pi / 6)\, d^3\, N\, a_{CM}(\lambda) \quad \text{.....4.3.3.2.4}$$

from which the chlorophyll specific absorption by equivalent solution $a_{SOL}^*(\lambda)$ can be obtained:

$$a_{SOL}^*(\lambda) = a_{SOL}(\lambda) / CHLa\ (ugL^{-1}) \quad \text{.....4.3.3.2.5}$$

This method for calculating $a_{SOL}^*(\lambda)$ makes assumptions regarding the shape of the cell, the absence of multiple scattering and thus the applicability of Beer's Law and is also reliant on measurement of cell volumes and cell numbers, it should therefore be only considered a useful approximation.

The $a_{SOL}^*(\lambda)$ for the phytoplankton used in the experimental tank are shown in Figure 21. The variation in strength of the package effect by wavelength is best described by the spectral flattening coefficient which is the ratio of $a_{PHY}(\lambda)$ to $a_{SOL}(\lambda)$. The spectral flattening coefficient is always less than 1 and tends to be closer to 1 when cells are small or absorption is weak, hence where the package effect is minimal (Sathyendranath *et al.* 1987). The spectral flattening for the phytoplankton used in the experimental tank are shown in Figure 22.

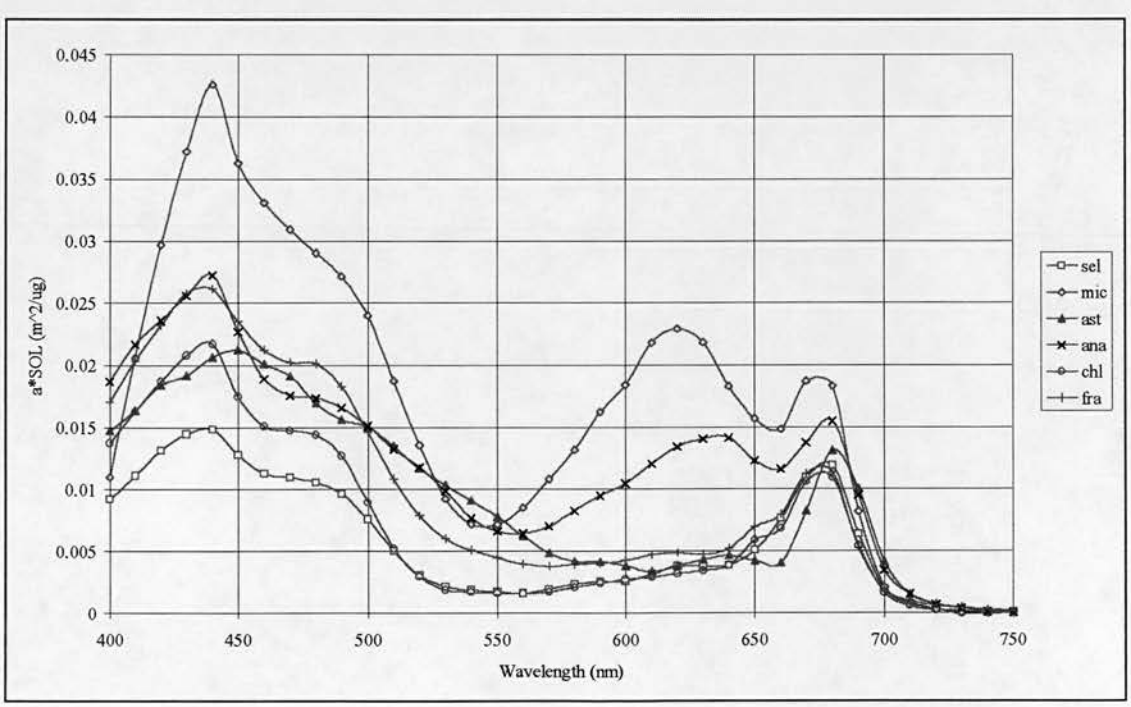

Figure 21. $a_{SOL}^*(\lambda)$ for the pure culture phytoplankton measured in the experimental tank. See Figure 20 for an expansion of legend.

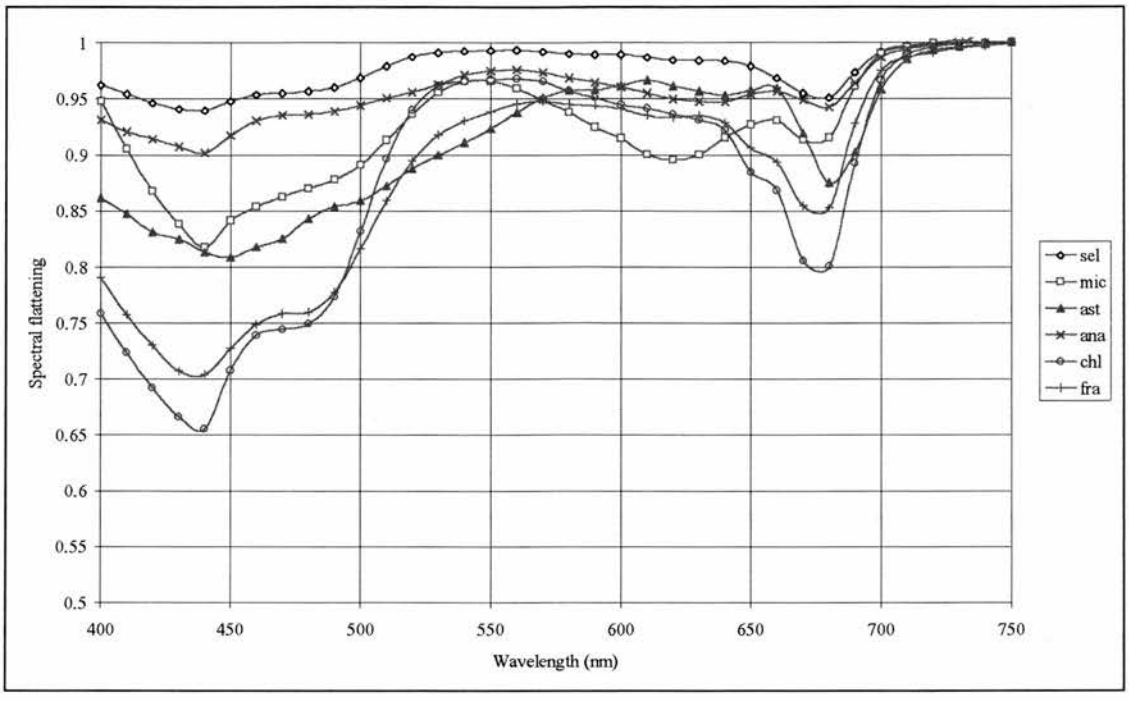

Figure 22. The spectral flattening of $a_{PHY}(\lambda)$. See Figure 20 for expansion of legend.

All the species had their lowest spectral flattening values at 440 nm and 675 nm which is where CHL*a* absorbs most strongly. These are the spectral regions where the package effect predicts the most significant flattening would occur. Similarly, the influence of the package effect was weakest where $a_{PHY}(\lambda)$ is minimal, which is beyond 700 nm. In that the strength of the package increases with smaller cell size, large cell numbers and higher intracellular pigment concentrations these parameters are presented in Table 6. On the basis of pigment concentration alone the data in Table 6 would suggest that *Asterionella* should experience the largest reduction in absorption efficiency because of its high intracellular pigment concentration, conversely *Microcystis* would experience the smallest reduction in absorption efficiency. However intracellular pigment concentration must be viewed in combination with cell size and cell numbers if the full influence of the package effect is to be calculated. At most wavelengths *Chlorella* experiences the most significant reduction in absorption efficiency due to the package effect, this is because it is the smallest cell and was at a high cell concentration, the influence of the package effect being slightly mitigated by the low intracellular pigment content. *Selenastrum* experienced the smallest reduction in absorption due to the package effect due to the lower cell numbers, increased cell size and low intracellular pigment content. The relatively large cells of the two Bacillariophyte species were present at low concentrations, however they both had high intracellular pigment concentrations consequently they both experienced significant reductions in absorption efficiency due to the package effect.

The importance of the package effect to this study is the influence that it has upon $R_{PHY}(\lambda)$. Those species with low spectral flattening values (equating with significant package effects) would require

larger increases in cell numbers to achieve the same increase in $a_{PHY}(\lambda)$ as those species with a minor package effect. Thus species with significant package effects will need more cells to register as a significant spectral influence on $R_{PHY}(\lambda)$ than would those species with less significant package effects. Another influence of the package effect is that increases in pigment concentration through increases in cell numbers or increases in intracellular pigment, both of which lead to increased $a_{PIG}(\lambda)$, will lead to diminishing increases in $a_{PHY}(\lambda)$ as the pigments compete for light thereby becoming individually less efficient absorbers. This latter effect has ramifications for CHL*a* retrieval algorithms based the reflectance trough associated with CHL*a* absorption.

Comparisons of $a_{SOL}^*(\lambda)$ have shown that although the package effect is an important factor in explaining variation in $a_{PHY}(\lambda)$ it could not account for all the variation with the remaining variation being attributed to differences in phytoplankton pigment composition.

4.3.3.3 The total absorption spectra

The reason for discussing the $a_{PHY}(\lambda)$ of the species used in the tank experiments is because of the inverse relationship between absorption and reflectance. The experimental tank set-up allows control of the optically important water quality parameters, however the phytoplankton are suspended in the sand-filtered water which is itself optically active consequently the spectral properties of $a_{H2O}(\lambda)$ must also be considered. For this study the temperature dependent $a_{H2O}(\lambda)$ measured by Hakvoort (1994) was used.

4.3.3.4 Comparison of scattering spectra for the phytoplankton used in the experimental tank

According to the modelling work of Morel and Prieur (1977) and Kirk (1981a, 1981b) reflectance is a function of the ratio of scattering to absorption, thus to fully understand $R_{PHY}(\lambda)$ it is important to measure not only the $a_{PHY}(\lambda)$ but also the phytoplankton scattering $b_{PHY}(\lambda)$. As mentioned in § 4.2.6 $b_{PHY}(\lambda)$ could not be measured given the equipment available however it could be approximated by the relative scattering spectra $\Delta\chi(\lambda)$ (Davis-Colley *et al.* 1986). To remove the effects of biomass concentration the CHL*a* specific relative scattering spectra $\Delta\chi^*(\lambda)$ can be used.

Figure 23 shows the $\Delta\chi^*(\lambda)$ for the phytoplankton species used in the tank experiments. The spectral form of these $\Delta\chi^*(\lambda)$ is very similar to the spectral form of the scattering efficiency spectra presented by Bricaud and Morel (1986) and Ahn *et al.* (1992). There are some spectral similarities between the species, notably the troughs in scattering at 440 and 675 nm. These troughs arise from the anomalous dispersion that occurs as the scattering efficiencies and absorption efficiencies interact and are centred upon the main absorption peaks. Consequently the specific $\Delta\chi^*(\lambda)$ for *Microcystis* and *Anabaena* show slight troughs at 620 nm. With the exception of *Microcystis* there are no major trends in spectral scattering throughout the visible wavelengths, *Microcystis* has a obvious inverse relationship between scattering and wavelength. The importance of this is that if there is a trending $\Delta\chi^*(\lambda)$ then it is likely

that there will be a trending $b_{bPHY}(\lambda)$, consequently more short-wavelength light will be reflected from the *Microcystis* cells than long-wavelength light. For species with trending $b_{PHY}(\lambda)$ the assumption that absorption is the inverse of reflectance does not hold. The factors influencing whether a species has a trending scattering spectra are morphological. Cells with dimensions similar to the wavelength of the interacting light and cells with gas vacuoles will both scatter more light and have inverse scattering relationship with wavelength (§ 2.1.5.2).

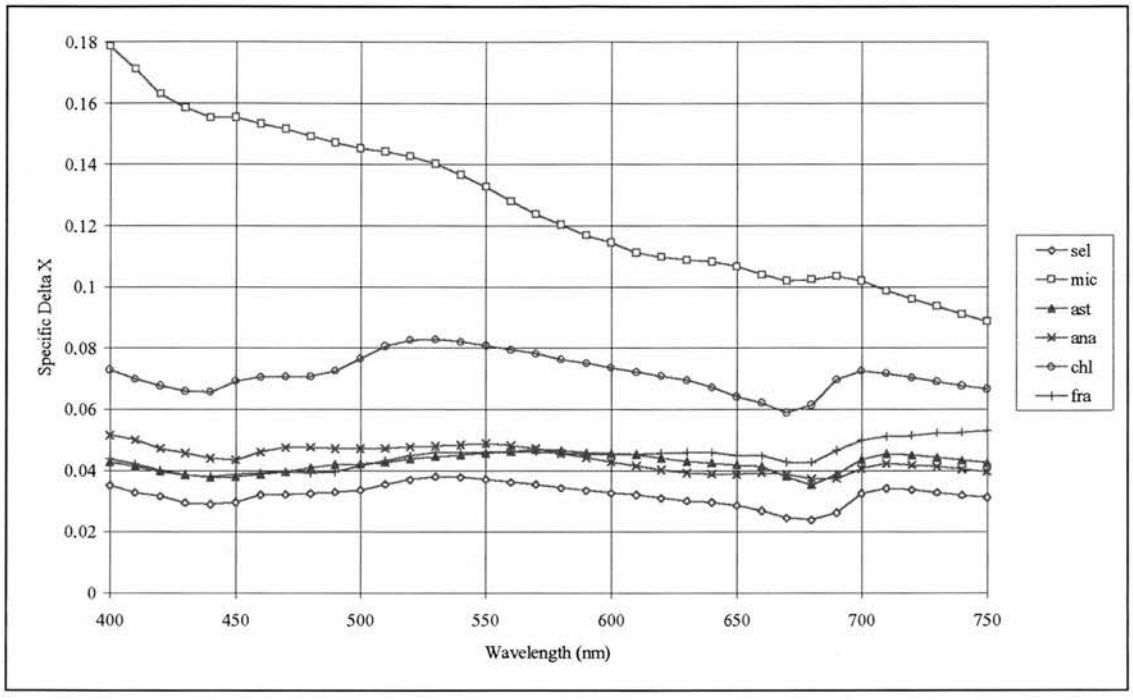

Figure 23. $\Delta\chi^*(\lambda)$ for the phytoplankton species used in the experimental tank. See Figure 20 for an expansion of the legend.

4.3.4 The pure culture phytoplankton reflectance spectra

The shape of spectral signature as measured in the experimental tank is assumed to be accurate although magnitude is not considered precise. This assumption is based upon the various factors concerning the experimental tank design and on the assumptions made when calculating the spectral signature. Notably the restriction placed upon photon pathlength by the 1.1 m tank depth, the shadow created by the tank side-walls, and the assumption concerning the Q factor all combine to cast uncertainty on the precision of the various spectral measurements required for the calculation of $R(\lambda)(-0.01\ m)$. These factors have no spectral dependency so the spectral nature of each measurement is assumed to be accurate, consequently the spectral signature may be assumed spectrally correct. This has implications for the nature of the conclusions that can be drawn from this study.

The individual species $R_{PHY}(\lambda)$ are presented below and their spectral properties are discussed in terms of within- and between-class differences. The biomasses for the nine concentrations of each

species measured in the tank are shown in Table 7. The individual species $R_{PHY}(\lambda)$ are presented below with reference made to the pigment features identified in the $a_{PHY}(\lambda)$ (§ 4.3.3).

Table 7. The CHL*a* ($\mu g\ L^{-1}$) concentrations for each species used in the experimental tank. Also shown is the concentration of the pure culture prior to dilution.

Series number	*Microcystis*	*Anabaena*	*Chlorella*	*Selenastrum*	*Asterionella*	*Fragilaria*
1	15.6	9.5	12.1	30.7	8.8	4.0
2	33.4	20.5	17.2	56.7	20.4	5.6
3	39.7	30.1	22.5	64.9	24.7	7.4
4	43.3	43.9	28.1	69.1	25.2	9.4
5	45.7	56.6	34.0	111.3	28.7	13.5
6	63.0	67.8	45.4	123.7	35.6	17.7
7	78.3	97.6	55.3	161.0	47.6	21.9
8	90.1	111.6	72.1	157.6	53.6	26.0
9	103.9	126.3	89.5	183.9	53.2	30.5
pure culture	1069.5	1201.5	819.4	1947.3	714.1	320.6

All the pure culture phytoplankton reflectance spectra show that below 450 nm $R_{PHY}(\lambda)$ is low. This is an artefact of the spectral properties of the light used to illuminate the experimental tank (Figure 15); at wavelengths shorter than 450 nm the flux output from the light is low hence the reflected flux is low which encroaches upon the signal-to-noise ratio of the SR9910 spectroradiometer.

The *Selenastrum* $R_{PHY}(\lambda)$ exhibited two peaks, a main peak at approximately 550 nm and a minor peak at approximately 700 nm (Figure 24). These peaks were interspersed with troughs at approximately 440, 675 and 740 nm with shoulders at 480 and 640 nm. These reflectance troughs correspond to absorption maxima for the pigments CHL*a* (the peaks at 440 and 675 nm) and CHL*b* (the shoulders at 480 and 640 nm). At wavelengths shorter than 510 nm lower concentrations reflected more light; beyond this inflection point the higher concentrations provided the strongest reflectance with the exception of the reflectance at 675 nm which was magnitudinally similar for all concentrations. Both the 550 nm and the near-IR peak were skewed.

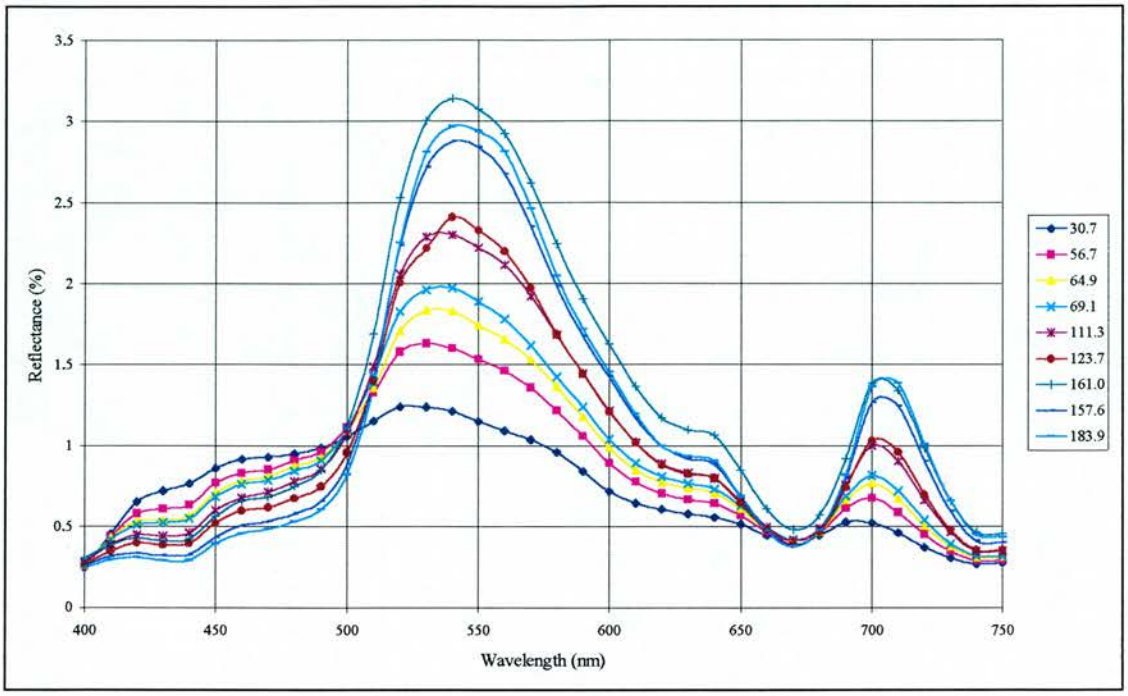

Figure 24. *Selenastrum* $R_{PHY}(\lambda)$ (%) with measured CHL*a* concentration (μg L^{-1}).

The *Chlorella* $R_{PHY}(\lambda)$ were essentially similar to the *Selenastrum* $R_{PHY}(\lambda)$ with the main reflectance troughs corresponding to those pigments mentioned above (Figure 25). The main peak was at 540 nm which was skewed towards the shorter wavelengths, there was a shoulder at 600 nm, a trough at 675 nm and a near-IR reflectance peak at 700 nm which was unskewed.

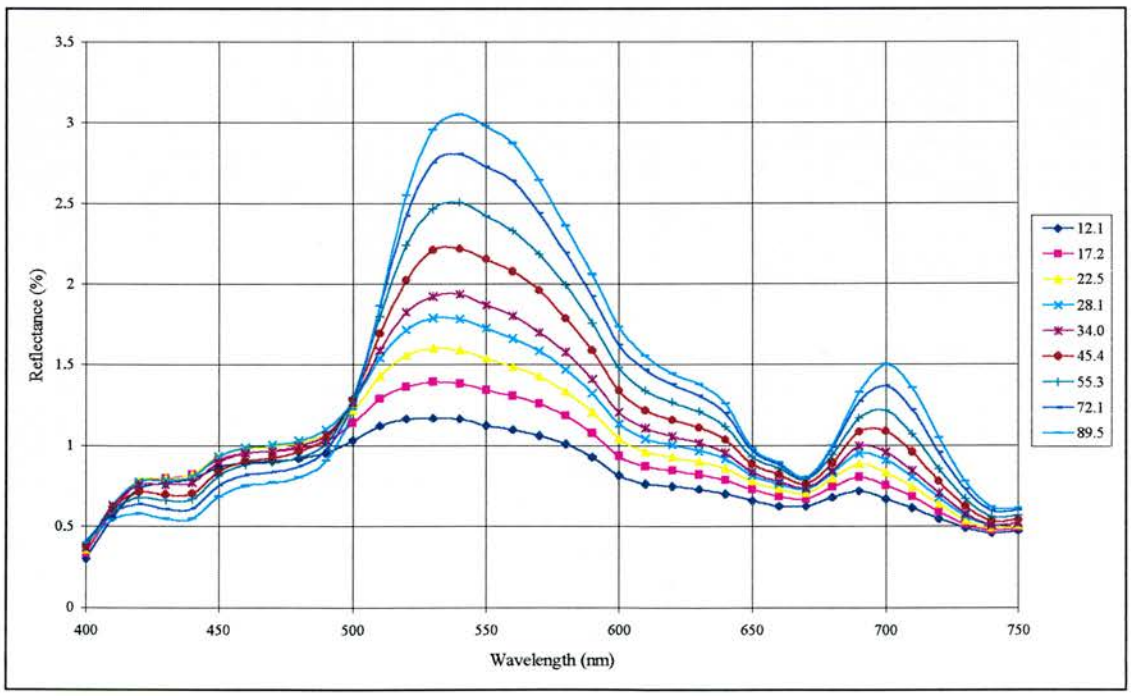

Figure 25. *Chlorella* $R_{PHY}(\lambda)$ (%) with measured CHL*a* concentration (μg L^{-1}).

There are more similarities between the two Chlorophyte spectra than there are differences. The *Chlorella* provided a larger $R_{PHY}(\lambda)$ per unit CHL*a*. The *Chlorella* had obvious shoulders at 600 and 650 nm while the *Selenastrum* spectra exhibited a minor feature at 625 nm.

Microcystis showed the tripartite peaks indicative of Cyanophytes with peaks at 540, 655 and 700 nm interspersed with troughs at 440 (CHL*a*), 620 (PCC), 675 (CHL*a*) and 740 nm (Figure 26). From the 540 nm reflectance peak to the far-red reflectance trough reflectance appears to become saturated above biomass concentrations of 40 mg CHLa L^{-1}. Above this concentration, increases in biomass do not result in increases in reflectance; this effect is not manifest in the 700 nm peak. This suggests that it would be difficult to obtain a reflectance from *Microcystis* of greater than 6%. From 440 to 500 nm reflectance increased slowly exhibiting subtle spectral features with the lower concentrations reflecting more than the higher biomass concentrations, from 500 to 540 nm reflectance increased dramatically and with the exception of the lowest concentration the spectra from the different concentrations conform following an inflection point at *ca.* 525 nm, this conformity continues from 500 to 675 nm. The local peak at 655 nm was skewed towards the longer wavelengths. In the near-IR wavelengths the highest concentrations reflected more than the lower concentrations.

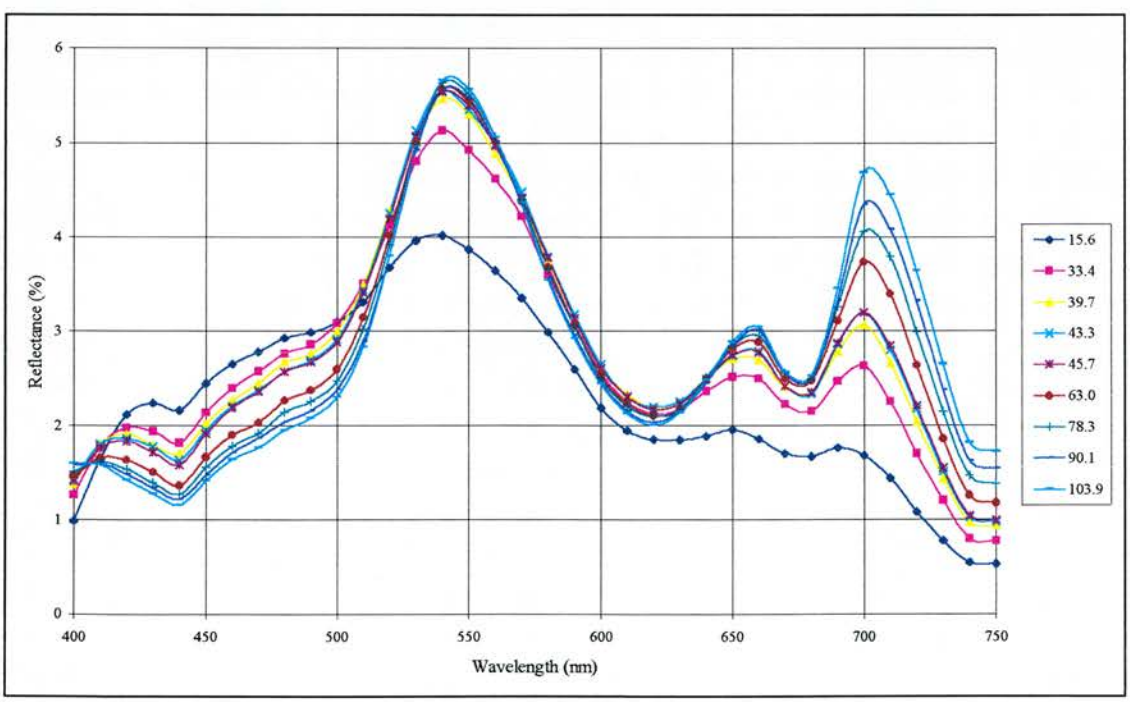

Figure 26. *Microcystis* $R_{PHY}(\lambda)$ (%) with measured CHL*a* concentration ($\mu g\ L^{-1}$).

The *Anabaena* $R_{PHY}(\lambda)$ also exhibited the tripartite peaks indicative of the Cyanophytes, peaking at 550, 655 and 700 nm, with troughs at 440, 620, 675 and 740 nm which are the locations of spectral features for those pigments described for *Microcystis*. The different concentrations exhibited

magnitudinal variation at most wavelengths and it is only at the highest measured biomass concentration that reflectance appears to show signs of saturation. At wavelengths shorter than the *ca.* 525 nm inflection point, the lower concentrations reflected more than the stronger concentrations, at wavelengths longer than 525 nm the strongest reflectance was generated by the higher biomass concentrations. There was a subtle trough at 510 nm at all concentrations.

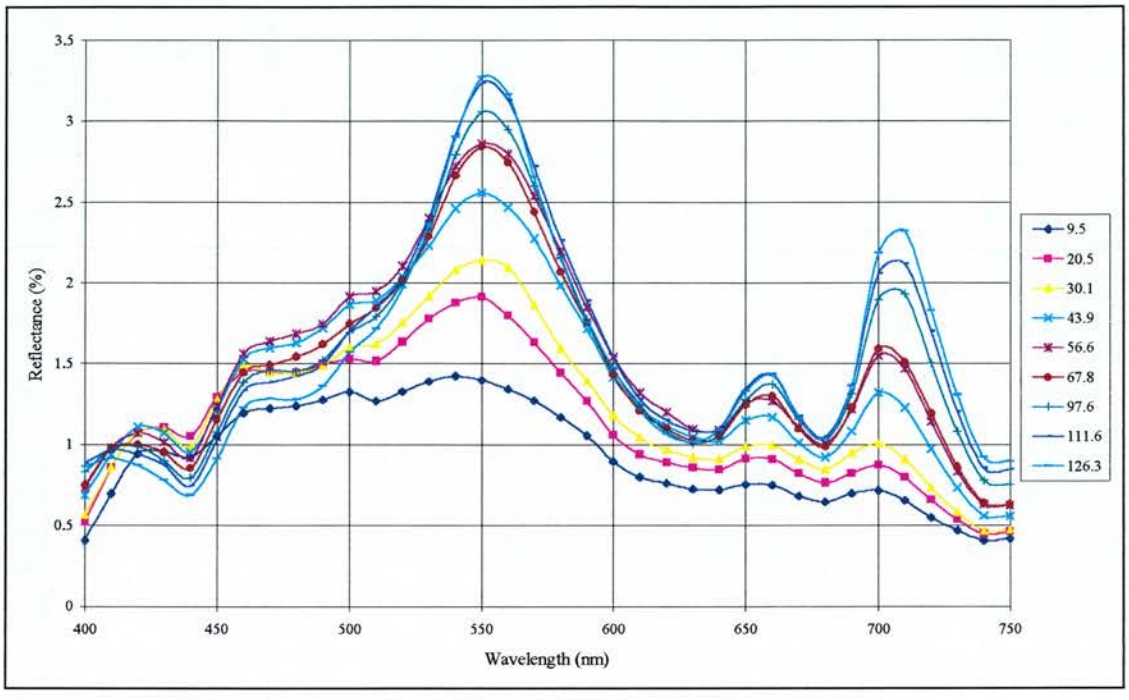

Figure 27. *Anabaena* $R_{PHY}(\lambda)$ (%) with measured CHL*a* concentration ($\mu g\ L^{-1}$).

The main difference between these two Cyanophyte $R_{PHY}(\lambda)$ was in the magnitude of reflectance, even at similar biomass concentrations the magnitude of the *Microcystis* reflectance was twice that of *Anabaena*. The spectral shape of the *Anabaena* $R_{PHY}(\lambda)$ is broadly similar to that of *Microcystis*, however there was a minor trough at 510 nm present in the *Anabaena* $R_{PHY}(\lambda)$.

Asterionella had a main reflectance peak at 565 nm and a lesser near-IR peak at 700 nm both of which were unskewed (Figure 28). There was also a minor peak at 650 nm. These peaks were interspersed with troughs at 440 (CHL*a*), 500 (FCX), 640 (CHL*c*) , 670 (CHL*a*) and 740 nm. Between 440 and 565 nm there were minor spectral features on the rising reflectance shoulder which were most likely attributable to CHL*c*. At wavelengths shorter than the *ca.* 500 nm inflection point the lower concentrations provided the strongest reflectance, at wavelengths above this inflection point reflectance was directly proportional to concentration. Typically the weaker concentrations exhibited less featured spectra than the stronger concentrations.

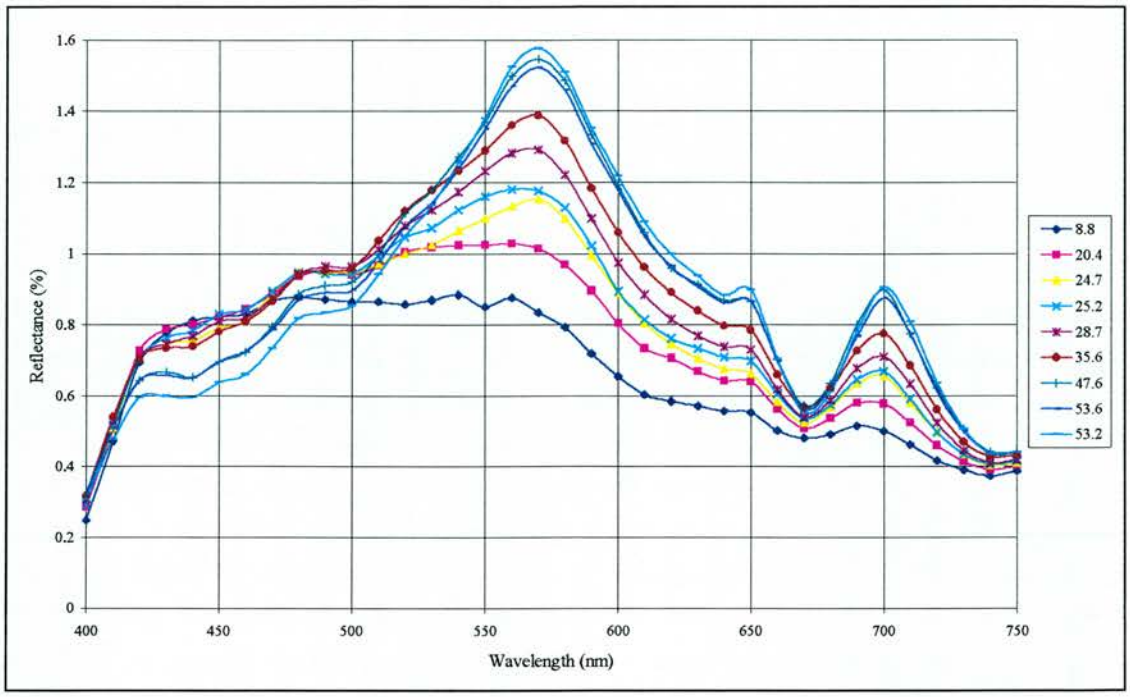

Figure 28. *Asterionella* series $R_{PHY}(\lambda)$ (%) with measured CHL*a* concentration (μg L^{-1}).

The *Fragilaria* $R_{PHY}(\lambda)$ was basically similar to the *Asterionella* $R_{PHY}(\lambda)$ with two main peaks at 565 and 700 nm, and troughs at 440 (CHL*a*), 600 (CHL*c*), 670 (CHL*a*) and 740 nm (Figure 29). Throughout the visible wavelengths reflectance magnitude is proportional to biomass concentration *i.e.* there is no inflection point in the *Fragilaria* concentration series $R_{PHY}(\lambda)$.

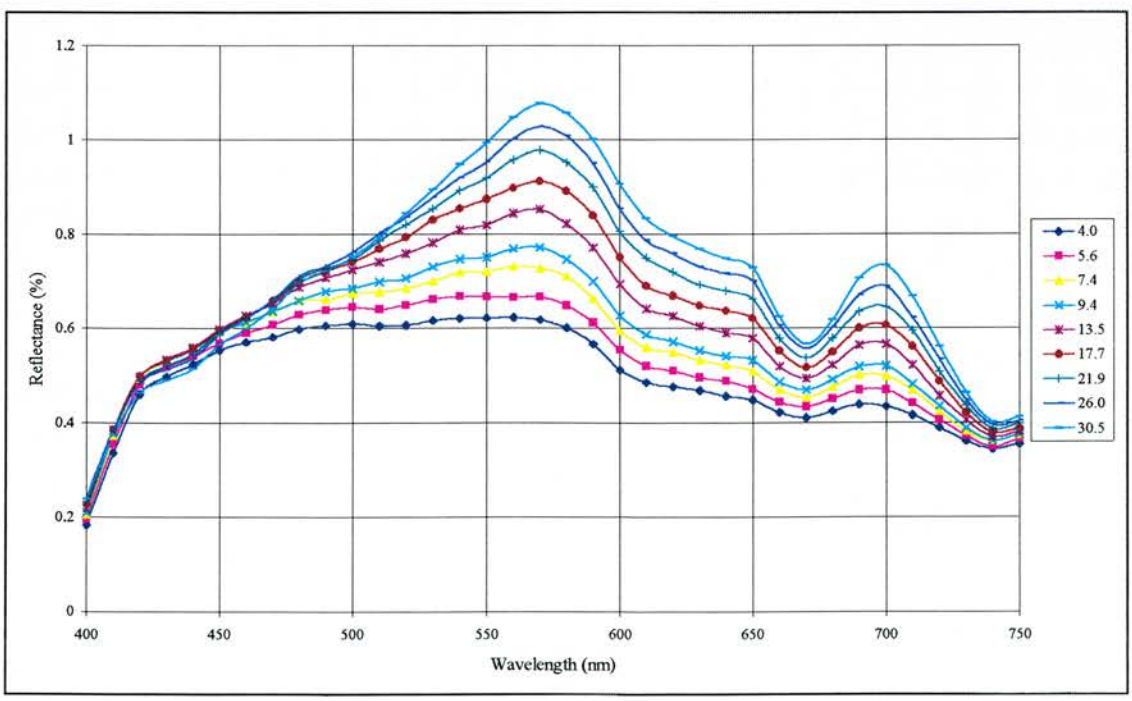

Figure 29. *Fragilaria* $R_{PHY}(\lambda)$ (%) with measured CHL*a* concentration (μg L^{-1}).

The *Asterionella* provides more per unit CHL*a* $R_{PHY}(\lambda)$ than did *Fragilaria*. Both the main *Fragilaria* peak and its near-IR peak were broader than the same peaks in the *Asterionella* spectra. The minor 650 nm reflectance peak of the *Asterionella* spectra was not present in the *Fragilaria* spectra.

The inverse relationship between the $a_{PHY}(\lambda)$ and $R_{PHY}(\lambda)$ is highlighted for *Anabaena* in Figure 30 which show the standardised (§ 3.1.3) absorption and reflectance spectra at the highest concentration in the measurement series. The $a_{PHY}(\lambda)$ peaks are accurately transposed to $R_{PHY}(\lambda)$ troughs. As $a_{PHY}(\lambda)$ peaks are absorption centres for pigments it is the trough locations in $R_{PHY}(\lambda)$ that are likely to be most useful in the identification of phytoplankton class from measurements of $R_{PHY}(\lambda)$. Note that the inverse relationship breaks down beyond 700 nm where $a_{H2O}(\lambda)$ significantly influences $R_{PHY}(\lambda)$ but is present in the phytoplankton absorption spectra.

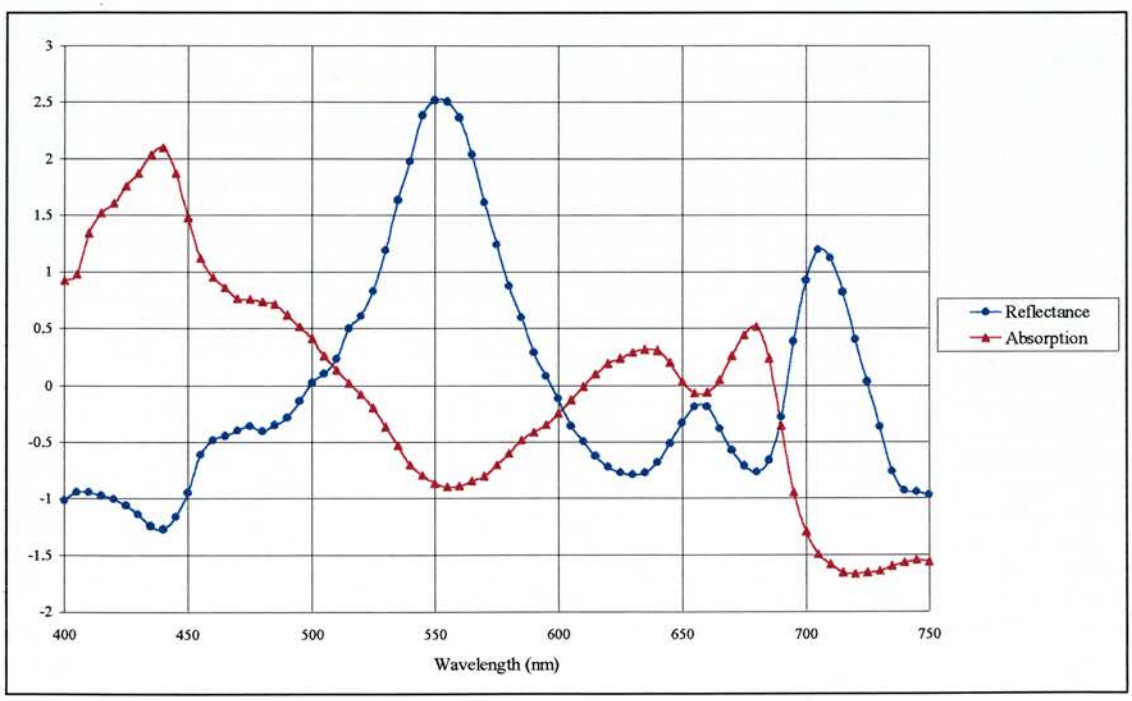

Figure 30. Standardised (§ 3.1.3) $a_{PHY}(\lambda)$ and $R_{PHY}(\lambda)$ for *Anabaena*.

The skews in $R_{PHY}(\lambda)$ arise from the spectral location and spectral width of the Gaussian shaped absorption curves which generate the reflectance troughs. The juxtaposition of the various pigment absorption bands and the water absorption features creates a total absorption spectra which is no longer a set of simple Gaussian features, this in turn creates skews in the reflectance spectra.

The pure culture $R_{PHY}(\lambda)$ presented here are spectrally similar to those presented by Quibell (1991, 1992). The reflectance spectra for a *Selenastrum* sp. at 650 μgL^{-1} CHL*a* showed no reflectance at wavelengths shorter than 475 nm (probably because of the tungsten-halogen light source), a 560 nm

peak, a 650 nm trough and a 725 nm peak (Quibell 1991). The peak at 725 μgL^{-1} was 1.4%. The above surface irradiance reflectance spectra for a *Microcystis* sp. at 124 μgL^{-1} CHL*a* showed the tripartite peaks typical of the Cyanophyceae, peaking at 550, 640 and 700 nm and troughing at 440, 610 and 675 nm (Quibell 1992). The peak at 550 nm reached 2%. For both species the near-IR peak exhibited the tendency to shift towards the longer wavelengths as biomass concentration increased.

4.3.5 Discussion of the pure culture reflectance spectra

To highlight the between-class spectral differences, Figure 31 shows $R_{PHY}(\lambda)$ at similar CHL*a* concentrations and Figure 32 shows standardised (§ 3.1.3) $R_{PHY}(\lambda)$ for the strongest CHL*a* concentrations.

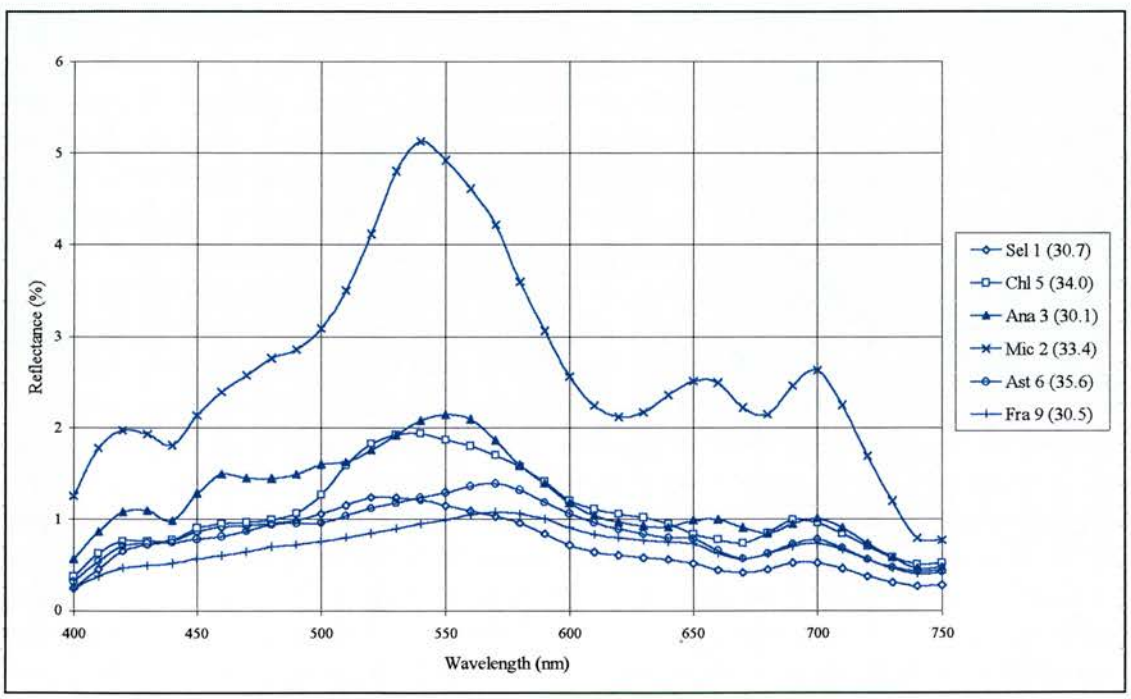

Figure 31. $R_{PHY}(\lambda)$ (%) for the individual species at similar CHL*a* concentrations ($\mu g\ L^{-1}$).

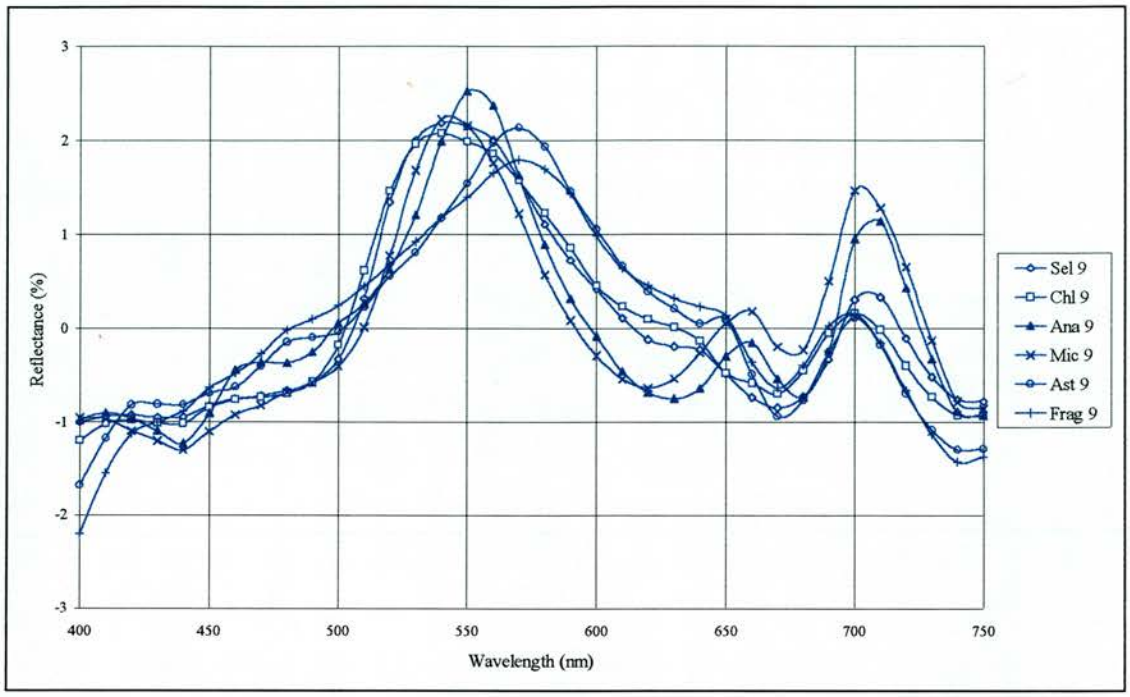

Figure 32. Standardised $R_{PHY}(\lambda)$ at the maximum CHL*a* concentration achieved for each series (see Table 7).

All the pure culture phytoplankton reflectance spectra exhibit broadly similar features highlighting the dominant influence of absorption by CHL*a* on $R_{PHY}(\lambda)$: a trough in the blue; a maxima in the green; a local trough in the far red; a reflectance peak in the near-IR and an unchanging spectra into the IR. Due to accessory pigmentary differences the locations of these features subtly differ between species and between classes. For wavelengths shorter than *ca.* 510 nm all species show reflectance to decrease with increasing biomass. This indicates that the increase in absorption by CHL*a* and other photosynthetic pigments at these shorter wavelengths is more significant than the increase in scattering. For wavelengths longer than *ca.* 510 nm all the species show directly proportional increases in reflectance with biomass concentration with the exception of *Microcystis* at its green peak. This indicates that the increases in absorption at these longer wavelengths are less significant than the increases in scattering, this holds true even at the far-red CHL*a* absorption peak. The apparent saturation of the *Microcystis* reflectance spectra from the green reflectance peak to the far-red reflectance trough suggests that the increases in scattering and absorption that occur with increases in biomass concentration are roughly equal. This results in no net change in reflectance with increasing biomass. Such a conclusion has implications for CHL*a* retrieval algorithms based on reflectance at these wavelengths.

The most apparent difference on a species level is the high $R_{PHY}(\lambda)$ from *Microcystis* (Figure 31). This arises because of the relatively high level of scattering by the small, spherical, gas vacuolate *Microcystis* cells (Figure 23). When individual species are considered there is a strong correlation

between CHL*a* concentration and $R_{PHY}(\lambda)$. At 750 nm, where $a_{PHY}(\lambda)$ is minimal, all variation in $R_{PHY}(\lambda)$ is due to changes in $b_{PHY}(\lambda)$ which increases proportionally to biomass concentration. This relationship is very species specific, when considering the spectra on a class level this relationship breaks-down (Table 8). This suggests that the relationship between $b_{PHY}(\lambda)$ and $R_{PHY}(\lambda)$ varies on a species level and not a class level with the exception of the Bacillariophyceae. The relatively large, silica covered *Asterionella* and *Fragilaria* cells are the least efficient scatters of the species considered (Figure 23). It may therefore be concluded that the magnitude of $R_{PHY}(\lambda)$ is not likely to be of use when trying to develop class identification routines.

Table 8. The correlation coefficients between $R_{PHY}(750)$ and CHL*a* concentration for both the individual phytoplankton species and classes.

Phytoplankton species or ***class***	*Correlation between $R_{(750)}$ and CHLa concentration*	*n*
Selenastrum	0.965	9
Chlorella	0.986	9
Chlorophytes	**-0.173**	**18**
Anabaena	0.996	9
Microcystis	0.997	9
Cyanophytes	**0.588**	**18**
Asterionella	0.958	9
Fragilaria	0.983	9
Bacillariophytes	**0.951**	**18**

The most striking between-class difference in the $R_{PHY}(\lambda)$ is that the Cyanophytes have three reflectance peaks with an extra peak at *ca.* 650 nm which is not present in the other classes (Figure 31 and Figure 32). This arises from the strong absorption at 620 nm by the biliprotiens C-PCC and C-PCE. On a more subtle level, the Bacillariophytes have a reflectance shoulder at 650 nm which is not present in the other classes, and the Chlorophytes have a reflectance shoulder at 640 nm which is also not present in the other classes.

There are between-class differences in the shape of the green (*ca.* 550 nm) reflectance peak for each class. The Chlorophytes have the steepest rise in reflectance combined with the shallowest fall resulting in peaks skewed towards the shorter wavelengths. The Cyanophytes have steep rising and falling shoulders resulting in the narrowest peaks. The Bacillariophytes have a combination of shallow rising and declining slopes resulting in the broadest peaks. Figure 31 shows that there is also a between-class difference in the location of the green reflectance peak which are summarised in Table 9. The Chlorophytes have their green peak located to the shortest wavelength (534 ± 1.19 nm), the Bacillariophyte green peak is at the longer wavelengths (567 ± 1.01 nm) and the Cyanophyte green reflectance peak lies between these two (545 ± 0.78 nm). Location of the green reflectance maximum can be used to model phytoplankton class with an F-ratio value of 261.1[4] which exceeds the

[4] The lowest concentration *Asterionella* $R_{PHY}(\lambda)$ did not exhibit a definite green peak so was excluded from this analysis.

critical F-ratio value at the 99.9% level of probability. With the exception of *Microcystis*, these locations are CHL*a* concentration dependent, with green peaks shifting towards the longer wavelengths as biomass increases (Table 9).

Table 9. The location of the green reflectance peak (*ca.* 550 nm) for the pure culture phytoplankton species and classes used in the experimental tank.

Phytoplankton *species* or **class**	*Average location of green reflectance peak (nm)*	*Correlation between location of the green reflectance peak and biomass*	*n*
Selenastrum	536 ± 2.04	0.817	9
Chlorella	533 ± 1.15	0.742	9
Chlorophytes	**534 ± 1.19**		**18**
Anabaena	548 ± 0.83	0.603	9
Microcystis	543 ± 0.50	0.360	9
Cyanophytes	**545 ± 0.78**		**18**
Asterionella	566 ± 0.50	0.554	9
Fragilaria	567 ± 1.31	0.771	9
Bacillariophyte	**567 ± 1.01**		**18**

The 440 nm reflectance trough is not always present at some of the lowest concentrations but when it is present its location does not vary between-classes consistently occurring at 436 ± 0.38 nm with a standard deviation of 2.27 nm. The Cyanophytes have a localised trough in the orange wavelengths, this occurs at all biomass concentrations at 628 ± 1.33 nm, varying with a standard deviation of 5.64 nm. There is a degree of between species variation in the location of this feature which occurs at 634 ± 0.88 nm for *Anabaena* and at 623 ± 0.38 nm for *Microcystis*.

All classes have a localised trough in the far-red wavelengths. Analysis of variance shows there to be a between-class variation in the location of this trough which can be modelled by class giving an F-ratio value of 131.32 which exceeds the F_{crit} of 3.17 at 99.9% probability level. The location of these troughs are given by class in Table 10. The Chlorophytes trough is at the shortest wavelengths, the Cyanophytes trough at the longest wavelengths and the Bacillariophytes trough falls between the two as can be seen in the Figure 32.

Table 10. The location of the far-red $R_{PHY}(\lambda)$ trough for each class.

Class	*Far-red trough location (nm)*	*n*
Chlorophyte	668.8 ± 0.29	18
Cyanophyte	678.6 ± 0.56	18
Bacillariophyte	670.5 ± 0.47	18

During the discussion of the $R_{PHY}(\lambda)$ for each species the trough features were cross-referenced against peak features in the $a_{PHY}(\lambda)$ as identified from the reported pigment absorption features compiled in the pigment database (Appendix Two). The troughs that are consistently present in the $R_{PHY}(\lambda)$ (*ca.* 440 and *ca.* 675 nm) are caused by the absorption by CHL*a* which was identified in the $a_{PHY}(\lambda)$ as being ubiquitous. Not all the pigment absorption features identified in the $a_{PHY}(\lambda)$ are visible in the

$R_{PHY}(\lambda)$. The 480 nm (CHL*b*) absorption for the Chlorophytes is present as a trough in the $R_{PHY}(\lambda)$ restricting the width of the green reflectance peak. There is no 635 nm (CHL*c*) feature however the 650 nm (CHL*b*) absorption feature is evident especially in the *Chlorella* $R_{PHY}(\lambda)$. The Cyanophyte $R_{PHY}(\lambda)$ show the *ca.* 625 nm trough (PCC) identified in the $a_{PHY}(\lambda)$. Similarly the 475 nm (carotenoid) and the 520 nm (PCC) absorption features noted in the *Anabaena* $a_{PHY}(\lambda)$ are present in the *Anabaena* $R_{PHY}(\lambda)$. However the 495, 570 and 585 nm (all PCE) absorption features are not evident in the $R_{PHY}(\lambda)$. For the Bacillariophyte $R_{PHY}(\lambda)$ the 460 nm (FCX, DDX or CHL*c*) and the 500 nm (FCX) absorption features of *Asterionella* are evident in the *Asterionella* $R_{PHY}(\lambda)$. Similarly the 480 nm (FCX) feature for the *Fragilaria* $R_{PHY}(\lambda)$ although the 620 nm (CHL*c*) feature is absent from both.

The apparent absence from the $R_{PHY}(\lambda)$ of some of the absorption features identified in the $a_{PHY}(\lambda)$ has implications for potential class identification routines unlike which are likely to rely upon assessment of pigment content through $R_{PHY}(\lambda)$. However more advanced pigment identification routines may highlight the more subtle pigment features that are not apparent from a simple visual assessment.

4.3.6 Summary of phytoplankton reflectance properties

This section has introduced the pure culture $R_{PHY}(\lambda)$ spectra which will be used in subsequent sections for the development and testing of class identification routines. The experimental tank in which these $R_{PHY}(\lambda)$ were measured has been fully described. The $a_{PHY}(\lambda)$ and $b_{PHY}(\lambda)$ spectra were also presented and their spectral influence on $R_{PHY}(\lambda)$ was discussed. Between-class differences and within-class similarities in $R_{PHY}(\lambda)$ have been identified, however the dominance of CHL*a* absorption features and the difficulty in identifying accessory pigment features in the $R_{PHY}(\lambda)$ has also been highlighted. The $R_{PHY}(\lambda)$ has been shown to be an inverse function of $a_{PHY}(\lambda)$ which in turn is a function of the intra-cellular pigment content of each species. Cell morphology does influence the magnitude of $R_{PHY}(\lambda)$ through the flattening of $a_{PHY}(\lambda)$ (the particle effect) and through $b_{PHY}(\lambda)$, however this influence is, in the first approximation, not spectrally trending.

Chapter Five Esthwaite Water

The controlled environment of the experimental tank does not fully represent the optical conditions the natural environment. In terms of both the radiative environment and the water quality the experimental tank was artificial. Reflectance spectra from a natural water body were therefore required for the purpose of testing the phytoplankton quantification and identification routines that are developed in the following two chapters. To this end repeated measurements of the reflectance spectra from Esthwaite Water were made. The reflectance spectra, along with accompanying absorption spectra, phytoplankton cell counts and water quality measurements, are presented in this chapter.

The first section in this chapter introduces Esthwaite Water and establishes the lake to be a Case I water body dominated by phytoplankton. The next section describes the various measurements that were performed. These are presented and discussed in terms of the phytoplankton population in the final two sections. The final section presents the surface and aerial reflectance spectra measured from Esthwaite Water.

5.1 Introduction to Esthwaite Water

Esthwaite Water lies in the south of the English Lake District at National Grid Reference SD358969. It is a small, relatively shallow lake with surface area of 1.01 km^2 (Figure 33). It has a complex shore line comprising of a number of discrete bays.

The production of phytoplankton in Esthwaite Water is considered to be phosphate limited. Study of the various phosphate sources led Hall *et al.* (1995) to the conclusion that natural phosphate loading from the surrounding catchment, combined with the residual loading from the sewage waste, waste inputs from a working fish farm and the phosphate reservoir in the sediments, had the capacity to maintain the phytoplankton community of Esthwaite Water between a mesotrophic or eutrophic status. The lake has been the subject of much study (Heaney *et al.* 1986 and the references therein).

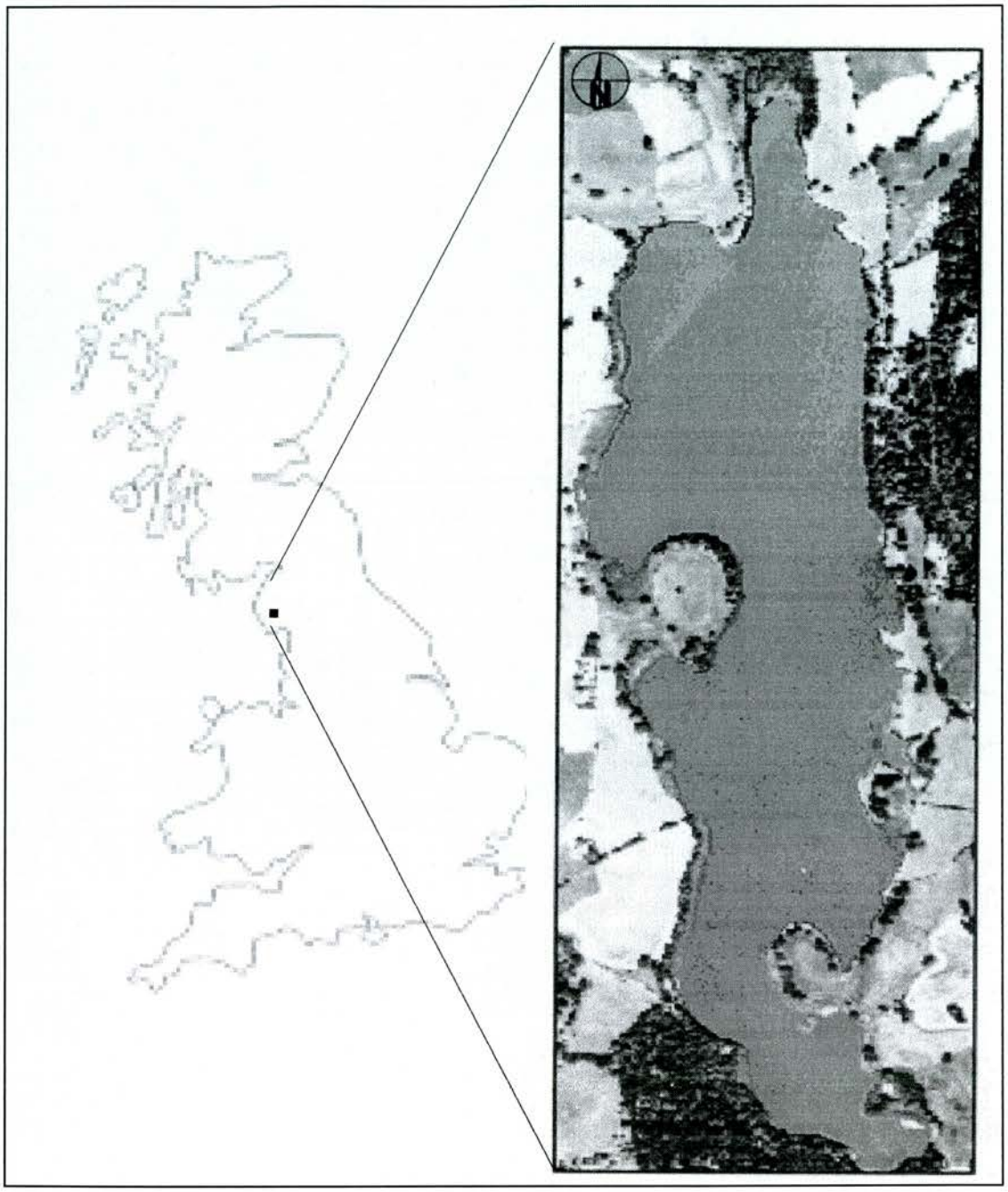

Figure 33. The location of Esthwaite Water within the UK. Also shown is an image of Esthwaite Water from the CASI flight acquired on 13/07/94 which highlights the many small bays in which phytoplankton blooms can accumulate due to wind induced movements.

Esthwaite Water was chosen for this study because its trophic status meant that it was capable of supporting a variety of phytoplankton crops at high densities throughout the year. Seasonal changes are manifest in both changes in the type of dominant species and changes in the overall phytoplankton biomass (Reynolds 1984, 1988). These periodic cycles of phytoplankton dominance can been summarised through the sequence of the identifiable dominant species beginning with a Bacillariophyte bloom in early spring, followed by a Chlorophyte bloom in late spring and early

summer and a Cyanophyte bloom for the remainder of the summer and early autumn (Reynolds 1988). Changes in species composition are the cumulative outcome of the response of the different species to seasonally recurrent environmental events. In particular the warming of the water body leads to thermal stratification which directly affects the nutrient supply and indirectly affects the light penetration to the epilimnon and hypolimnion (in a thermally stratified lake nutrient availability is linked directly to phytoplankton growth and phytoplankton biomass directly affects the attenuation coefficients). The seasonal activities of planktonic grazers also play an important role in the seasonal fluctuations of phytoplanktonic succession. Superimposed upon these longer term environmental changes are short timescale changes which even on a small lake can impose a spatial pattern on the density of a phytoplankton crop. Localised environmental influences (Figure 33) cause spatial variation in nutrients through a combination of complex bathymetry, point nutrient input sources, spasmodic mixing from the littoral to deep water zones by thermal plumes, wind-induced water movement and the Coriolis force (George 1993). These spatially localised environmental influences, coupled with conveyor-belt concentration of buoyant gas-vacuolate phytoplankton due to wind mixing (George 1993), lead to the creation of 'patches' of relatively high concentrations of phytoplankton in the surface waters of Esthwaite Water.

Although a single species may be dominant in the waters of Esthwaite Water there is frequently a coincidence of many sub-dominant species. Seldom does the domination by one species lead to the exclusion of others (Reynolds 1988). This coexistence of species is explained by the ability of phytoplankton species to occupy various specialised niches and the greater the number of niches the greater the diversity of the coexisting species (Tilman *et al.* 1982).

A study of the underwater light climate of lakes in the English Lake District, which included Esthwaite Water, was made by Pearsall and Hewitt (1934) and Talling (1971) who concentrated on the attenuation of light in respects to the overall productivity. A time-series of Esthwaite Water reflectance spectra was measured in a study of how the optical properties of the Lake District lakes varied with water quality parameters (Charlton 1990). However no attempt was made to develop algorithms or phytoplankton identification routines. CHL*a* retrieval algorithms for surface and airborne reflectance spectra from Esthwaite Water were developed by George and Charlton (1996).

5.2 Water quality and radiometric sampling methods

Esthwaite Water was sampled on 16 occasions: six times in 1994; nine in 1995; and once in 1996. Over this period sampling was restricted to late spring, summer and early autumn. On each occasion the sub-surface irradiance reflectance $R(\lambda)(-0.01\ m)$ was measured. The method employed for measuring $R(\lambda)(-0.01\ m)$ during 1994 differed from that employed during 1995 and 1996. To more fully understand the changes in reflectance concomitant analysis of the absorption spectra of the seston and dissolved fractions of the water body was undertaken. CHL*a* and phytoplankton cell counts

were also made. The water samples on which these analyses were performed were taken from the surface layer. Samples were taken in 2 L pre-washed polypropylene bottles. If the laboratory analysis was not performed on the same day as the radiometric analysis then the samples were stored in a darkened 4°C cold room. Samples for cell counts were fixed with Lugol's iodine and the counts performed according to method described in § 4.3.2. Samples for DOC analysis were immediately filtered through a 0.2 μm average pore size GF/C filter and transferred into acid-washed glass bottles and stored in a refrigerator until the DOC analysis was conducted. The concentration of DOC was measured according to Tipping *et al.* (1988) using a Phase Separation Ltd. TOCsinII analyser running in continuous mode with the limit of detection set to 0.5 mg C L^{-1}. The dry weight per unit volume of a sample was measured according to Jones (1979) and ashed dry weight according to Standard Methods (1992).

Secchi disk depth Z_{SD} (m) has traditionally been used as an assessment of water clarity and has often been used to provide an estimate of downwelling attenuation coefficient for the PAR wavelengths (Edmondson 1980, Priesendorfer 1986, Effler *et al.* 1991). Z_{SD} (m) was recorded at each sample station with a 0.3 m diameter flat white disk lowered to depth until is disappeared from view, whereupon it was raised until it reappeared and the depth below the surface recorded.

5.2.1 Radiometric measurements

The 1994 radiometric measurements were performed *in situ*. These field measurements were taken from a boat anchored in an area of the water body deep enough to avoid bottom reflectance. The measurements were performed with a Spectron SE-590 spectroradiometer[5] (Spectron Engineering, Colorado). The SE-590 has been used by many workers in the field of aquatic remote sensing including Maracci and Ooms (1988), George and Hewitt (1990) and Matthews and Boxall (1994). The SE-590 has a photodiode array to measure radiance in 252 spectral channels over a range of the electromagnetic spectrum from 400 to 1100 nm. Measurements were made at bandwidths of 2.6 to 2.8 nm, with a spectral resolution of approximately 13 nm. The instrument automatically measures a dark current reading before each scan and sets a suitable integration time (between $16/60^{th}$ and $64/60^{th}$ of a second). Eight scans were made, which were automatically averaged and stored on magnetic tape. A 15° FOV optic was used for all spectral measurements. The scans were downloaded to a PC and calibrated against pre-measured calibration files. The calibration of the SE-590 was performed against a lamp calibrated by the National Physics Laboratory. The protocol for measuring reflectance from water bodies was discussed in § 2.3.1.

The 1995 spectroradiometric measurements were made in the experimental tank on 70 L samples collected from the surface of Esthwaite Water. $R(\lambda)$ was measured in the experimental tank on the

[5] On loan from the Environmental Science Department, University of Nottingham, Sutton Bonnington Loughborough, Leicestershire UK.

same day as the samples were collected. The protocol for measuring R(λ) in the experimental tank with the Macam SR-9010 is described in § 4.3.1.2.

5.3 Variation in the quality of Esthwaite Water

Changes in phytoplankton crop are discussed in terms of biomass and population dynamics. Variation in other optically active water quality parameters during the sampling period are also discussed and evidence for the assumption that Esthwaite Water being a Case I water body is given.

5.3.1 Variation in the phytoplankton crop

To give an indication of the phytoplankton biomass CHL*a* was measured concomitantly with the radiometric sampling using the hot methanol extraction method described in § 6.1. During the 1994 sampling season CHL*a* concentrations were initially low, increased until mid-summer, then slowly decreased towards autumn (Table 11). During the 1995 season concentrations began much lower and stayed relatively low until 25/07/95 when there was an eight fold rise to 127.9 μgL^{-1}. Concentrations then dropped back down to 23.6 μgL^{-1} two weeks later, rose steadily again peaking at the end of summer and dropping in the autumn. The 1995 sequence exhibits a mid-summer clear water period up until 25/07/95 when a patchy bloom occupied the sample site.

Table 11. The summary phytoplankton population at the time of the radiometric sampling. See text for explanation of species diversity index

Sample date	*CHLa* (μgL^{-1})	*Dominant species*	*Dominant class*	*Total cells /mL*	*Species diversity index*
18/05/94	22.9	No sample	No sample	No sample	No sample
06/07/94	56.7	*Anabaena*	Cyanophyte	6230	0.69
14/07/94	59.3	*Oscillatoria*	Cyanophyte	11236	0.75
28/07/94	55.6	*Anabaena*	Cyanophyte	16299	0.31
01/09/94	48.0	*Anabaena*	Cyanophyte	12804	1.16
21/09/94	37.8	*Anabaena*	Cyanophyte	11521	0.96
23/05/95	7.2	*Sphaerocystis*	Chlorophyte	3196	0.5
06/06/95	19.9	*Cryptomonas*	Cryptophyte	1963	0.53
20/06/95	16.1	*Rhodomonas*	Cryptophyte	4506	1.43
04/07/95	16.4	*Anabaena*	Cyanophyte	7862	1.11
25/07/95	128.0	*Anabaena*	Cyanophyte	105604	0.61
01/08/95	23.9	*Anabaena*	Cyanophyte	15822	1.14
15/08/95	38.1	*Oscillatoria*	Cyanophyte	21259	1.1
12/09/95	45.6	*Oscillatoria*	Cyanophyte	11630	0.96
10/10/95	25.5	*Anabaena*	Cyanophyte	602	1.56
10/04/96	12.9	*Asterionella*	Bacillariophyte	3710	0.85

A more complete picture of the variations in phytoplankton biomass in Esthwaite Water is provided by the fortnightly CHL*a* measurements performed as part of the Institute of Freshwater Ecology (IFE) monitoring programme (Figure 34). The spring bloom, the mid-summer clear water, the full summer bloom and the winter decline in phytoplankton biomass are highlighted by Figure 34.

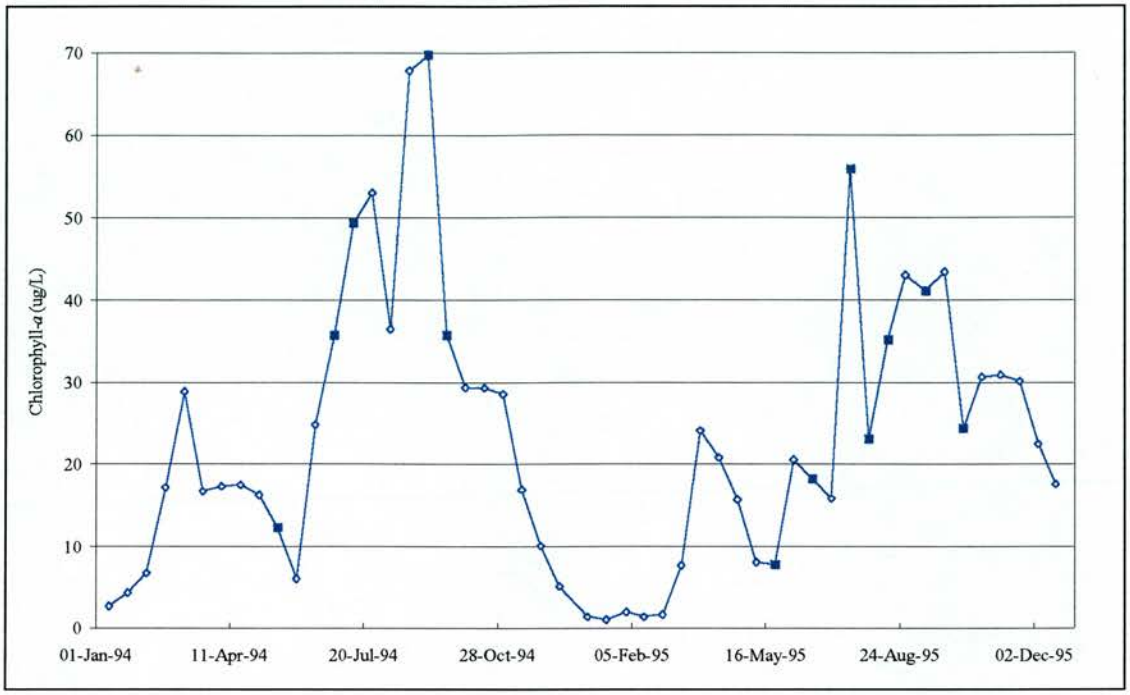

Figure 34. CHL*a* ($\mu g\ L^{-1}$) concentrations measured as part of the IFE Esthwaite Water monitoring programme. Filled square data-points are those which were approximately coincident with the sampling visits of this study.

Surface water samples for phytoplankton cell counts were taken contemporaneously with the radiometric samples. The results of this analysis are summarised in Table 11 and Figure 35. Table 11 gives the dominant species, dominant class, total number of cells, species diversity index along with CHL*a* concentration. The species diversity d, at the time of sampling is an index which relates the number of species present to the total number of (Reynolds 1988). d is calculated according to:

$$d = (s - 1)/\ln(N) \quad \text{.....5.3.1.1}$$

where s is the number of species present at a concentration > 0.1 mL^{-1} and N the total number of cells. The larger the d value the more diverse the species composition. Figure 35 shows the variation in the two most prolific species in each class over the two years of the radiometric sampling.

The correlation coefficients between total cell numbers and the CHL*a* concentration and dry weight concentrations were both strongly positive with an R of 0.868 and 0.977 respectively. This suggests that CHL*a* concentration was a fair approximation for phytoplankton biomass and that the major fraction of the seston component of the water body was phytoplankton.

The contemporaneous cell counts were performed upon surface water samples. However some phytoplankton species can regulate their position within the water column and others have a tendency to sink out of the water column under calm conditions. Species may also accumulate in enclosed bays

due to wind-induced movements. Thus spot surface water samples are unlikely to fully represent the phytoplankton crop both spatially and throughout the water column. However they should represent the relative phytoplankton assemblages in the water sample viewed by field spectroradiometers.

A better synopsis of the annual changes in the phytoplankton crop in Esthwaite Water is provided by the IFE fortnightly sampling. This survey includes five metre depth integrated samples for cell counts. Sampling the top five meters of the water body means a wider range of species are encountered when compared to surface water samples which are frequently biased towards those species capable of self regulating their buoyancy. The five metre sampling provides a better overview of the lakes planktonic community while the surface water sample is a more accurate assessment of those plankton responsible for generating the reflectance signal.

The fortnightly sampling of Esthwaite Water showed that the calendar year 1994 was generally warmer and wetter than average thereby provided more favourable conditions for the development of blooms, blue-green scums and anoxic events (Reynolds *et al.* 1995). These conditions were offset by the weather variability which, through frequent disturbances of the successional progress, has been shown to be a key factor in maintaining low average planktonic biomass (Reynolds 1993). The spring diatom bloom reached a maximum of 28.8 μg CHL*a* L^{-1} in March, with Bacillariophytes continuing to dominate being limited only by silica availability until the onset of stratification in early summer when they sank through the thermocline. There then followed a period of diversity with a mixture of

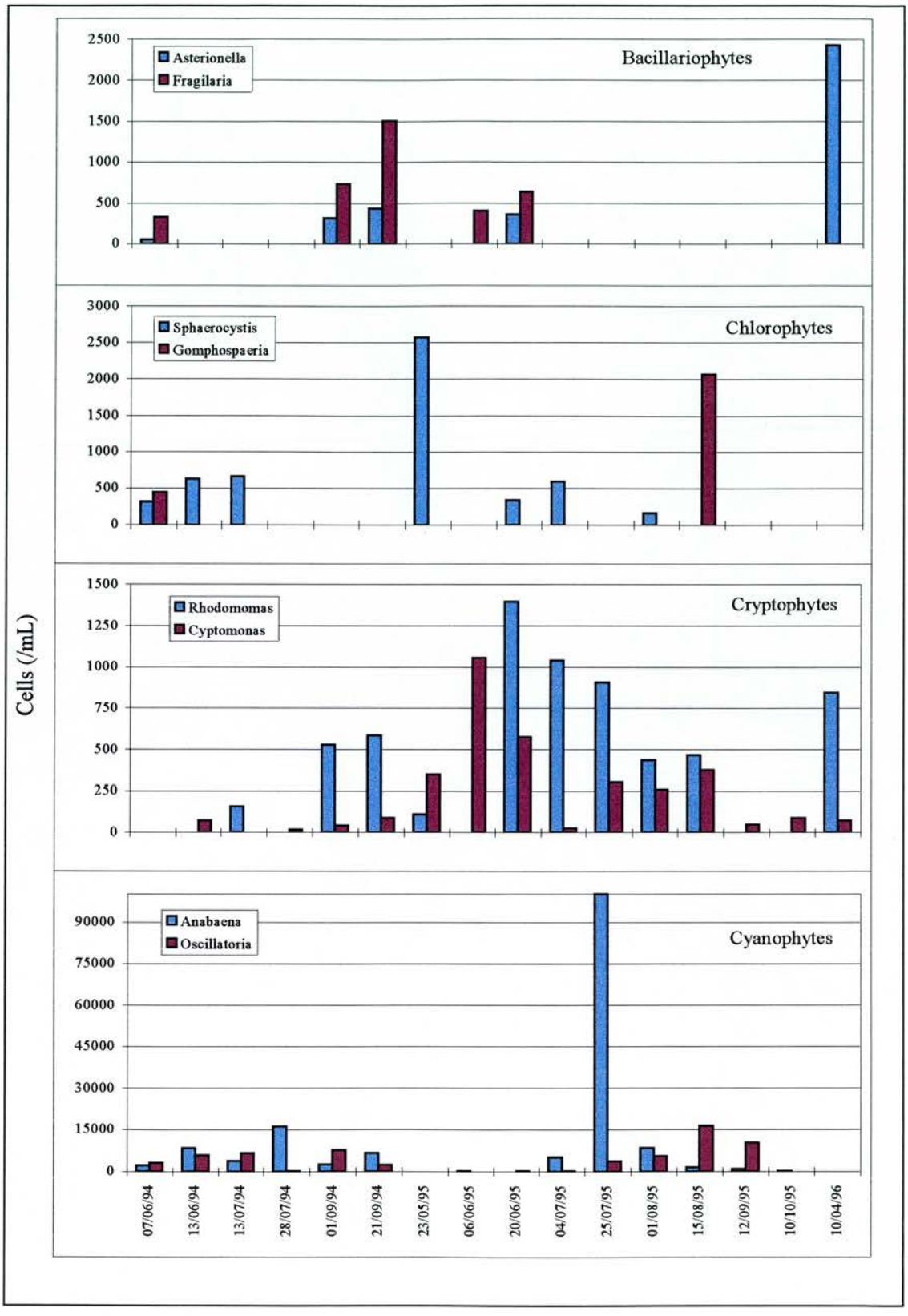

Figure 35. The two most prolific phytoplankton species from the four main classes found in the Esthwaite Water surface water samples taken contemporaneously with the radiometric samples (Bacillariophytes, Chlorophytes, Cryptophytes and Cyanophytes). Note the vertical axes for each class differ and the horizontal axis is not date-proportional.

Chlorophytes and Cyanophytes, as the latter reached a maximum restricted by nitrogen and phosphorus availability. 1995 had a very wet winter followed by a warm dry summer (Reynolds *et al.* 1996). The spring Bacillariophyte bloom peaked in late March, the onset of stratification was followed by the ascendancy of Cryptophytes until severe phosphate depletion characterised the system and Cyanophytes bloomed in both July and late August.

5.3.2 Esthwaite Water clarity and quality

Z_{SD} measurement was performed to provide an indication of the optical quality of Esthwaite Water at the time of each radiometric survey. Z_{SD} results are given in Table 12. Relatively deep Z_{SD} readings indicate that light attenuation is low and that the water is clear. Shallow Z_{SD} readings such strong light attenuation associated with poor water clarity.

Z_{SD} was found to be inversely correlated with CHL*a* with a correlation coefficient of -0.569. This correlation suggest that the optical clarity of Esthwaite Water is primarily driven by its phytoplankton component with optical clarity decreasing as phytoplankton biomass increases. This is confirmed by Figure 36 which shows a strong relationship between optical clarity and phytoplankton biomass as recorded during the IFE fortnightly sampling of Esthwaite Water.

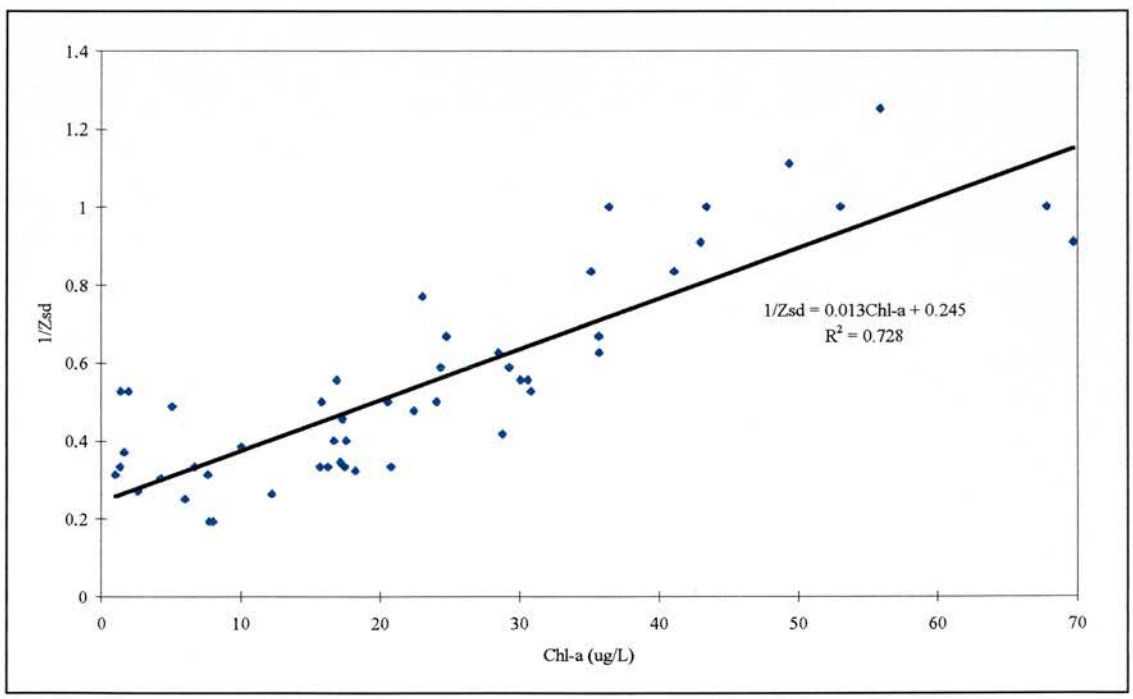

Figure 36. The regression analysis for Esthwaite Water CHL*a* and Z_{SD} confirming the relationship between optical quality and phytoplankton biomass. ($1/Z_{SD}$ is used solely to give a positive relationship). Measurements of both parameters were performed by IFE staff during 1994 and 1995.

DOC can be coloured or non-coloured. Coloured DOC absorbs light in the shorter wavelengths, whilst non-coloured DOC can be no affect upon the optical characteristics of the water body (§ 2.1.4). The

optical properties of DOC are described by the g_{440} parameter which is the absorption by the dissolved fraction at 440 nm. Concentrations of DOC in Esthwaite Water were typically within the range experienced by Tipping *et al.* (1988) who found that summer DOC values in the lake exceeded winter values and suggested a link between DOC and biological activity. There was no correlation between DOC and g_{440} (R = 0.194) hence there was no relationship between DOC concentration and DOC colour. There was only a weak correlation between CHL*a* concentration and DOC.

Table 12. The CHL*a*, Secchi depth Z_{SD}, dry weight, Dissolved Organic Carbon DOC, g_{440} (§ 2.1.4) and the parameters describing the $a_{DYS}(\lambda)$ at the time of the radiometric surveys (Equation 2.1.4.1).

Sample date	*CHLa* (μgL^{-1})	Z_{SD} *(m)*	*Dry weight* (μgL^{-1})	*DOC* (mgL^{-1})	g_{440} (m^{-1})	*ln* $a_{(DYS)}$ *slope*	*ln* $a_{(DYS)}$ *intercept*
18/05/94	22.9	2.25	1.5	2.0	1.04	-0.0124	5.49
14/07/94	59.3	1.25	3.7	3.7	2.14	-0.0099	4.71
28/07/94	55.6	0.95	2.69	2.69	0.85	-0.1201	5.11
01/09/94	48.0	1.25	3.33	3.7	0.98	-0.0138	6.10
21/09/94	37.8	1.25	3.3	2.69	1.43	-0.0132	6.33
23/05/95	7.2	4.75	1.4	3.86	0.68	-0.0122	5.03
06/06/95	19.9	1.67	2.67	2.42	0.99	-0.0111	4.92
20/06/95	16.1	2.79	0.7	3.52	1.04	-0.0094	4.23
04/07/95	16.4	2.5	2.1	2.2	1.05	-0.0102	4.59
25/07/95	128.0	1.18	180	3.1	1.43	-0.0094	4.54
01/08/95	23.9	1.3	4.8	6.69	1.51	-0.0089	4.37
15/08/95	38.1	1.3	6.8	3.01	0.95	-0.0104	4.58
12/09/95	45.6	0.85	6.66	2.72	0.88	-0.0105	4.51
10/10/95	25.5	1.7	4.5	2.85	1.25	-0.0119	5.48
10/04/96	12.94	2.7	1.8	5.01	0.95	-0.0104	4.48

Measured dry weight correlates strongly with CHL*a* concentration giving a correlation coefficient of 0.827. Dry weight and total cell numbers also correlated well. These correlations provide further evidence for the hypothesis that Esthwaite Water is a Case I water body. This is further confirmed by the evidence from the 1995 sampling season when ashed dry weight measurements were performed. These showed that ashed dry weight was typically 16.7% of total dry weight. The ashed fraction of dry weight is the suspended mineral particle contribution of tripton to seston which in Case I waters is low.

5.4 Esthwaite Water absorption spectra

Throughout this study the strong inverse relationship between the inherent optical property absorption and apparent optical property reflectance has been emphasised. On the basis of this relationship the radiometric analysis was complemented with laboratory-based measurements of the absorption properties of Esthwaite Water. An understanding of the total absorption spectra should lead to a fuller understanding of the reflectance spectra from Esthwaite Water. The ability to identify pigment absorption features in seston absorption spectra facilitates the identification of pigments in the reflectance spectra. As discussed in § 2.1.5 seston absorption (a_{SES}), which includes absorption by

phytoplankton and tripton (a_{PHY} and a_{TRI}), is combined with DYS absorption and water absorption (a_{DYS} and a_{H2O}) to give total absorption (a_{TOT}). The combined effects of $a_{H2O}(\lambda)$ and a strong $a_{DYS}(\lambda)$ can restrict the spectral observation window to between 500 and 700 nm. It is within this window that $a_{PHY}(\lambda)$ is strongest.

5.4.1 Esthwaite Water DYS absorption spectra

Measurement of $a_{DYS}(\lambda)$ has been performed by Kirk and Tyler (1986), Davis-Colley and Vant (1986), Malthus and Dekker (1990) for the purposes of interpreting the underwater light climate, and by Tipping *et al.* (1988) for the quantification of dissolved organic matter in lakes in the English Lake District. For this study all $a_{DYS}(\lambda)$ measurements were performed using a method adapted from Davis-Colley and Vant (1986). Field samples were filtered through a GF/C filter into acid-washed glass bottles on site. Samples were then chilled and stored in the dark prior to measurement. Measurement between 400 and 750 nm was performed using 100 mm optically matched glass cuvettes in a Hitachi U2000 spectrophotometer described in § 4.2.3.1. Three replicate spectra were measured and the average calculated.

The g_{440} values indicate considerable variation in $a_{DYS}(\lambda)$ in Esthwaite Water (Table 12). Over the sample period g_{440} values began the year low, rose through the summer and fell in the autumn. The $a_{DYS}(\lambda)$ spectra for Esthwaite Water are presented in Figure 37. The exponential relationship between a_{DYS} and λ is clearly visible. There were samples when g_{440} was obviously high *e.g.* from 14/07/94 to 28/07/94 $a_{DYS}(\lambda)$ rose from 0.99 to 2.14 m^{-1}. A g_{440} value of 2.14 m^{-1} represents a considerable absorption sink for photons at the shorter wavelengths. Also shown in Table 12 are the slope and intercept parameters for the $\ln(a_{DYS})(\lambda)$ which can be used to describe the exponential form of $a_{DYS}(\lambda)$ over the visible wavelengths when the natural log of a_{DYS} is regressed against λ (Equation 2.1.4.1). The intercept parameter is the hypothetical $\ln(a_{DYS})$ at 0 nm. The slope parameter falls within the range of published slope values collated by Dekker (1993).

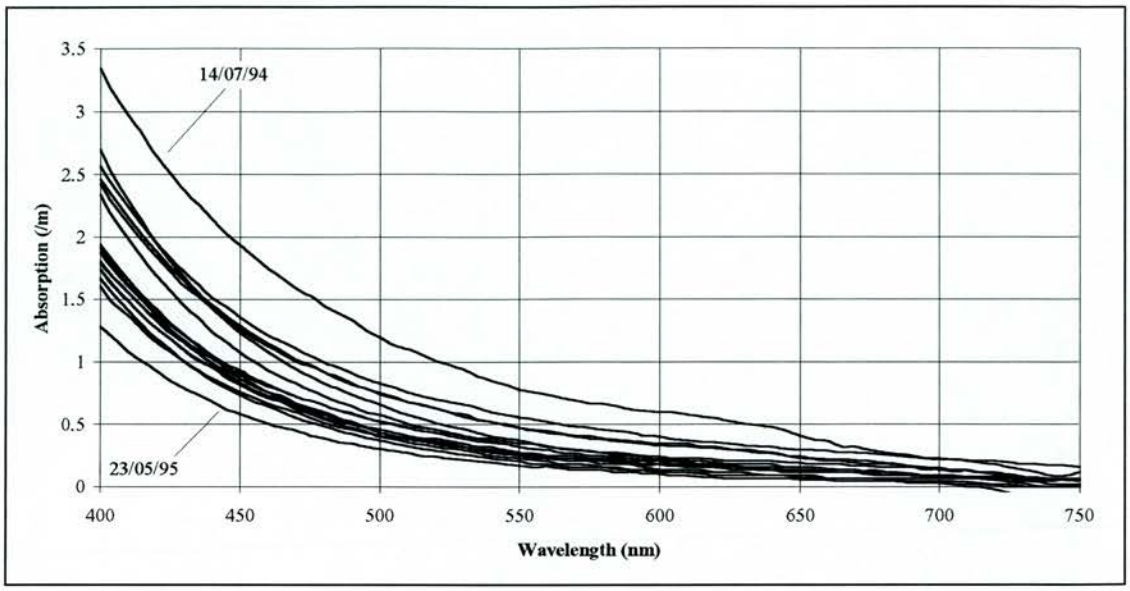

Figure 37. $a_{DYS}(\lambda)$ for Esthwaite Water. The strongest and weakest $a_{DYS}(\lambda)$ have been identified, the other samples can be identified through their g_{440} measurements (Table 12).

5.4.2 Esthwaite Water seston absorption spectra

The $a_{SES}(\lambda)$ were measured according to the modified version of the method of Davis-Colley *et al.* (1986) (§ 4.2.3.1). Field samples from freshwaters with low seston concentrations required concentration to achieve a reasonable signal-to-noise ratio. The sample concentration method was adapted from the methods of Kirk (1994) and Malthus and Dekker (1990). A known volume of the sample was vacuum filtered onto a 1.0 μm pore size Sartorious membrane filter pad. Retained particles were then resuspended in 3 mL of distilled water by a combination of back-flushing with air for 1 minute and gently rubbing the surface of the filter pad with a small sable-hair paint brush for a further 3 minutes. The filter pad was then rinsed in 2 mL of distilled water. $a_{SES}(\lambda)$ required correction for this concentration after the measurements were performed. Checking the filter pad under a microscope after resuspension suggested that most of the retained particles were resuspended. However the degree of resuspension success for each sample was not known. This represents a source of error that could lead to an underestimation of the magnitude of $a_{SES}(\lambda)$. No undertaking was made to separate the tripton signal from the phytoplankton signal.

The $a_{SES}(\lambda)$ spectra have been converted into $a_{SES}^*(\lambda)$ according to Equation 2.1.5.2.2 to remove the influence of phytoplankton biomass and allow comparisons of the spectral differences in attributable to pigment absorption features. The Esthwaite Water $a_{SES}^*(\lambda)$ are shown in Figure 38. The spectra have been grouped by the dominant phytoplankton class identified in the cell counts at the time of sampling. Figure 38a and b are Cyanophyte dominated spectra and Figure 38c are spectra dominated by phytoplankton from other classes. All the $a_{SES}^*(\lambda)$ exhibit two principal absorption peaks, the Soret absorption band at 440 nm and the far-red absorption band at 675 nm. Both of these are due to

absorption by the ubiquitous pigment CHL*a*. Troughs in the $a_{SES}^*(\lambda)$ are consistently present at 400 and 650 nm where there is limited pigment absorption. There are no $a_{SES}^*(\lambda)$ features at wavelengths beyond 710 nm because phytoplankton photosynthetic pigments do not absorb in the near-IR and there was little tripton present.

The Cyanophyte dominated $a_{SES}^*(\lambda)$ (Figure 38a and b) have an absorption peak at 625 nm which is not present in the other spectra. The 625 nm absorption peak is caused by the presence of biliproteins, and principally by the presence of PCC which is present in all Cyanophytes. The presence of this absorption feature creates an absorption trough at 600 nm. These spectra also show a small absorption peak at 575 nm which is also indicative of biliproteins, but in this instance it is PCE. The 25/07/95 $a_{SES}^*(\lambda)$ has a double peaked Soret absorption band, peaking at both 440 and 480 nm. The 480 nm peak in the 25/07/95 spectra is probably attributable to carotenoids.

The $a_{SES}^*(\lambda)$ dominated by non-Cyanophyte classes (Figure 38c) exhibit no strong 625 nm absorption feature, and they have fewer features between 440 and 650 nm. There are no distinguishing features visible to the eye which would facilitate the identification of the dominant phytoplankton class for these $a_{SES}(\lambda)$, all that could be concluded is that, by the absence of the 625 nm peak, they are not dominated by Cyanophytes.

5.4.3 Esthwaite Water total absorption spectra

The $a_{TOT}(\lambda)$ (Figure 39) show how the other optically active components combine to mask some of the features of the $a_{SES}(\lambda)$ (Figure 38). With the exception of 25/07/95 the $a_{TOT}(\lambda)$ spectra between 400 and 550 nm are dominated by $a_{DYS}(\lambda)$. However some of the $a_{SES}(\lambda)$ features do remain visible as subtle shoulders between 400 and 550 nm. Between 550 to 700 nm $a_{SES}(\lambda)$ features were visible in the $a_{TOT}(\lambda)$, notably the absorption at 625 and 675 nm. The 25/07/95 spectrum was measured when CHL*a* concentration was high (127.9 μg CHL*a* L^{-1}); at these concentrations of phytoplankton $a_{PHY}(\lambda)$ becomes considerably more significant than $a_{DYS}(\lambda)$. This is apparent in the total absorption spectra for 25/07/95 which exhibits an absorption peak at 440 nm and spectral features on the slope from 440 to 550 nm. In all the spectra wavelengths above 700 nm are dominated by $a_{H2O}(\lambda)$. $a_{H2O}(\lambda)$ is essentially constant and in the absence of tripton $a_{TOT}(\lambda)$ show little variation beyond 710 nm. Certain $a_{H2O}(\lambda)$ features are visible in $a_{TOT}(\lambda)$ as shoulders at 600 and 660 nm. This is particularly true at low phytoplankton concentrations.

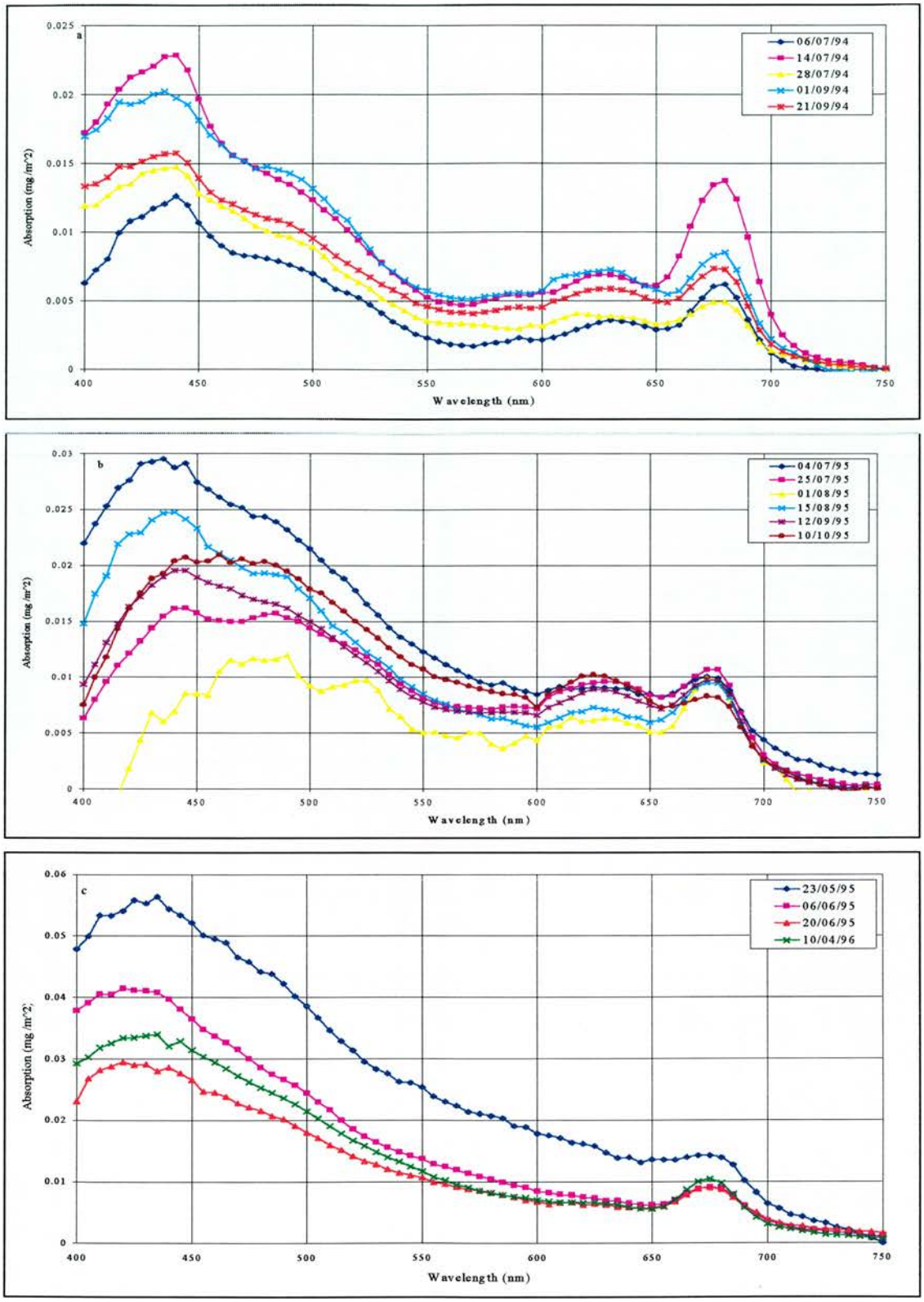

Figure 38. The $a_{SES}^*(\lambda)$ (m^2mg^{-1}) for Esthwaite Water. The spectra have been grouped according to dominant class type as identified by cell counts. The spectra in Figure 38a and b were dominated by Cyanophytes. The spectra in Figure 38c were dominated by a Bacillariophyte (10/4/96), a Chlorophyte (23/5/95) and two Cryptophytes (6/6/95 and 20/6/95).

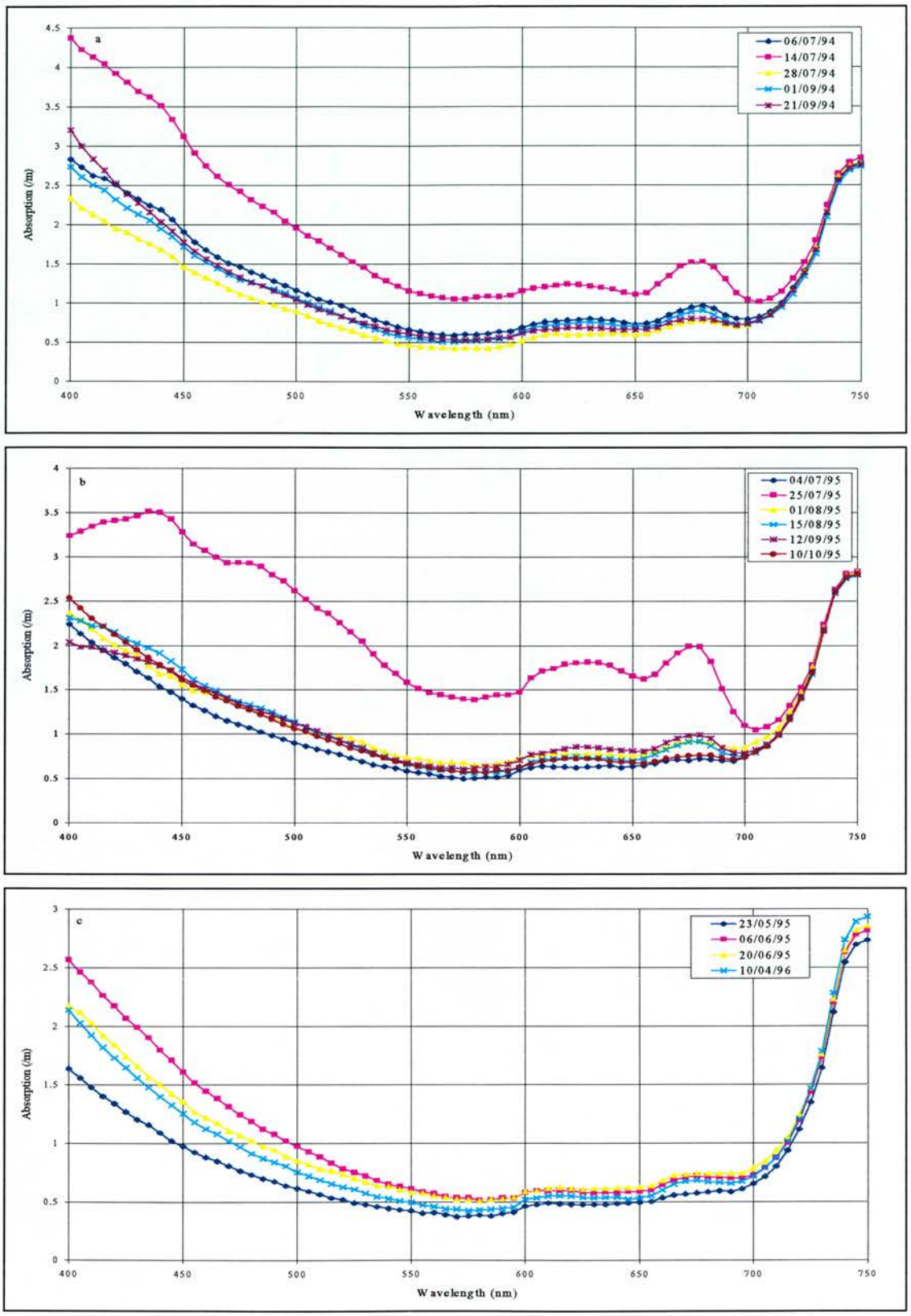

Figure 39. The $a_{TOT}(\lambda)$ for Esthwaite Water. The spectra have been grouped according to dominant class type as identified by cell counts. The spectra in Figure 39a and b were dominated by Cyanophytes. The spectra in Figure 39c were dominated by a Bacillariophyte (10/4/96), a Chlorophyte (23/5/95) and two Cryptophytes (6/6/95 and 20/6/95).

5.5 Esthwaite Water reflectance spectra

Sixteen spectral signatures were measured for Esthwaite Water between 1994 and 1996. Six were performed in the field and ten in the experimental tank. Sub-surface irradiance reflectance R(-0.01 m)(λ) was calculated on each occasion because it is independent of surface and atmospheric effects that can be present in other forms of the spectral signature and is therefore most suitable for multi-temporal comparisons (Dekker 1993).

5.5.1 Esthwaite Water sub-surface irradiance reflectance spectra

The R(-0.01 m)(λ) spectra are presented in Figure 40a, b and c. The spectra have been grouped in the same way as the $a_{SES}^*(\lambda)$ in § 6.4.2. For display purposes the spectra in Figure 40 have been smoothed using a 3rd order polynomial, 7 point Savitzky-Golay smoothing filter (§ 3.1.1).

The R(-0.01 m)(λ) have basically similar spectral properties. The principal reflectance peak occurs at *ca.* 560 nm, with a secondary peak at *ca.* 700 nm. Some of the spectra have a third peak at 655 nm. Reflectance peaks are separated by troughs at *ca.* 440, 625, 675 and 740 nm. Superimposed upon these main features are minor troughs and peaks which are more aptly described as shoulders. The higher the phytoplankton biomass concentration the more prominent these peaks become.

Assuming that spectral variation in scattering is not significant then the reflectance features are principally generated by features in the total absorption spectra. Peaks in $a_{TOT}(\lambda)$ can be directly related to troughs in reflectance and as these $a_{TOT}(\lambda)$ peaks have been shown to be pigment absorption features then reflectance troughs are related to pigment composition. The most clearly defined reflectance trough in Figure 40 is at 675 nm which is the centre of one of CHL*a* absorption bands. CHL*a* is ubiquitous in phytoplankton so this feature is present in all reflectance spectra where phytoplankton are present at moderate to high concentrations. The Cyanophyte dominated reflectance spectra have a reflectance trough at 625 nm caused by the presence of PCC. The presence of both these pigments was identified in the $a_{TOT}(\lambda)$. However the R(-0.01 m)(λ) exhibit few of the spectral features identified between 400 and 550 nm in the $a_{SES}^*(\lambda)$ due to the masking effect of $a_{DYS}(\lambda)$ at these wavelengths. The 400 to 550 nm $a_{SES}(\lambda)$ features have been reduced to shoulders on the reflectance spectra. It is between 550 and 710 nm that $a_{SES}(\lambda)$ features dominate the R(λ)(-0.01 m).

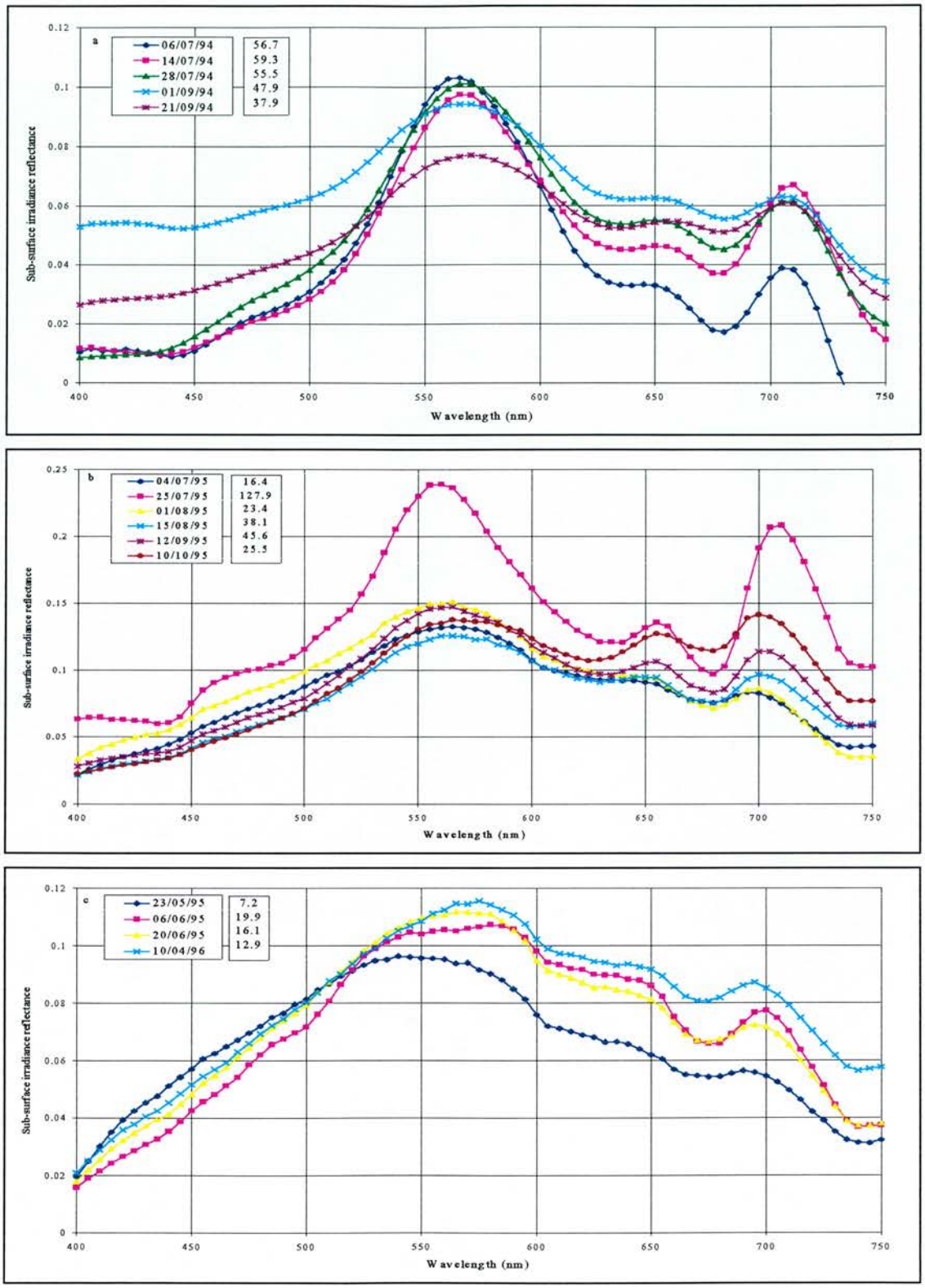

Figure 40. The measured R(-0.01m)(λ) for Esthwaite Water. The spectra have been grouped according to dominant class type as identified by cell counts. The spectra in Figure 40a and b were dominated by Cyanophytes. The spectra in Figure 40c were dominated by a Bacillariophyte (10/4/96), a Chlorophyte (23/5/95) and two Cryptophytes (6/6/95 and 20/6/95). The CHL*a* concentrations (μgL^{-1}) are shown adjacent to the legend entry. Note the differences in *y*-axis scale.

The spectra in Figure 40a were Cyanophyte dominated. *Anabaena* was the dominant species on 6/7/94, 28/7/94, 1/9/94 and 21/9/94, and *Oscillatoria* on 14/7/94. The species diversity during the *Oscillatoria* dominance was higher than during the periods of *Anabaena* dominance. All these spectra exhibit the three peaked spectral signature indicative of Cyanophytes which is caused by the presence of PCC. The spectra in Figure 40b and Figure 40c were dominated by Chlorophytes (23/5/95), Cryptophytes (6/6/95 and 20/6/95), Bacillariophytes (10/4/96) and Cyanophytes (4/7/95, 25/7/95, 1/8/95, 15/8/95, 12/9/95 and 10/10/95). The Cyanophyte dominated spectra in Figure 40b also exhibit the three peaked spectral signature characteristic of this class. This latter sample had a relatively low dominance combined with a high species diversity. The spectra dominated by non-Cyanophyte phytoplankton (Figure 40c) exhibit only two peaks.

The most obvious between-class difference in these lake reflectance spectra are the three peaks with a primary peak at 560 nm which is indicative of Cyanophytes. Two peaks with a primary peak located at 560 nm is indicative of Bacillariophytes. Two peaks with a primary peak located at 575 and 540 nm is indicative of Cryptophytes and Chlorophytes dominated waters respectively.

The primary peak location for the Bacillariophyte dominated spectra was at 575 nm, the Chlorophyte dominated spectra had a primary peak location at 540 nm, and the two Cryptophyte dominated spectra had primary peak location at 570 and 580 nm. The variation in primary peak location is determined by the location of the $a_{TOT}(\lambda)$ minima which is a function of pigment composition. There were no other reflectance features visible to the eye which are indicative of phytoplankton class.

5.5.2 Esthwaite Water CASI hyperspectral reflectance spectra

To provide further spectra on which to test the quantitative and qualitative algorithms developed during this study some Esthwaite Water reflectance spectra were obtained at altitude. As part of a collaborative monitoring exercise between the National Rivers Authority (NRA) and the IFE a study of the lakes in the English Lake District was conducted (George and Charlton 1996). Included in this project were some flights of the Compact Airborne Spectrographic Imager (CASI).

The CASI functions both as a multispectral imager (spectral mode) and as an imaging spectrometer (spatial mode). In spectral mode the device has a spectral resolution of 2.6 nm from 430 to 950 nm. In spatial mode the CASI is unique in respect to other imaging spectrometers in that the number of bands and width of these bands are programmable. In spectral mode the bands can not overlap and are restricted to a maximum of 15. Increasing the number of bands is at the expense of the spatial resolution (Zacharias *et al* 1992).

The CASI was flown above Esthwaite Water at an altitude of 1000 m on 13/07/94. The CASI was operated firstly in spatial mode and then in hyperspectral mode. The images and spectra were

analysed using Erdas Imagine v8.2 running on a Sparc 20 UNIX workstation. Six profiles of the CASI hyperspectral above surface upwelling radiance were sampled each being an average of 15 pixels. These $Lu_a(\lambda)$ spectra were splined to 2 nm intervals and atmospherically corrected using reference measurements made with the Spectron SE-590 at ground level. They were converted into above surface radiance reflectance by ratioing the CASI $Lu_a(\lambda)$ against $Ed_a(\lambda)$ as measured by the SE-590 at the water's surface using a Lambertian reference panel. These reflectance spectra are shown in Figure 41. No attempt was made to correct for the diffuse component of downwelling irradiance specularly reflected from the water's surface which forms a component of $Lu_a(\lambda)$ (Dekker 1993).

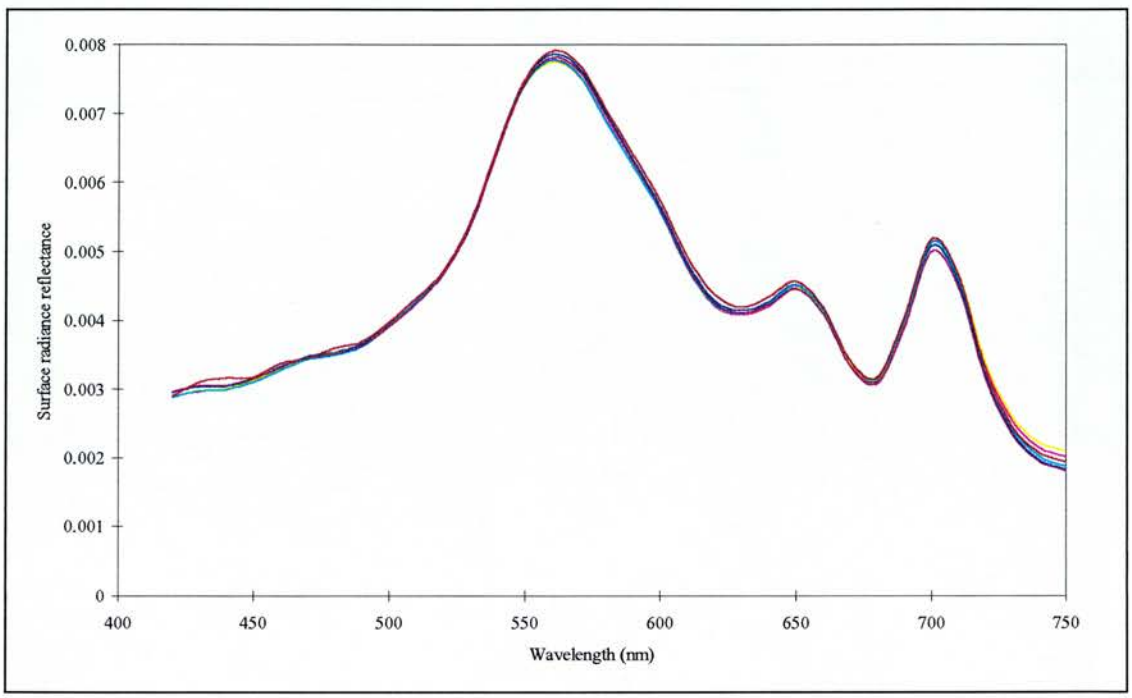

Figure 41. Six CASI hyperspectral reflectance spectra of Esthwaite Water measured at 1000 m on 13/07/94.

Cell counts of the actual phytoplankton assemblages were performed upon samples taken on the same day as the CASI flight from sample sites in close proximity to the sites from which the spectral profiles were obtained. These counts were conducted as part of the NRA's on-going survey of Esthwaite Water.

5.6 Concluding remarks for the Esthwaite Water

Esthwaite Water is a Case I water body with the main determinant of optical quality being phytoplankton biomass. The optical clarity of the lake varies with phytoplankton biomass with light being unable to penetrate to depth during the summer months due to the attenuating properties of the phytoplankton.

Esthwaite Water has a dynamic phytoplankton population being subject to periodic phytoplankton blooms throughout the spring, summer and autumn. During the summer months the lake was dominated by Cyanophytes while Bacillariophytes and Chlorophytes bloomed in early spring then remained consistently present at low levels through-out the year. Cryptophytes remained present through-out the year. Phytoplankton were the main component of the seston fraction of the water body. DOC concentrations (measurement of which is described in § 5.2) were high (a mean of 2.86 mg L^{-1} for Esthwaite during 1994 compared to a mean for the lakes sampled through-out the English Lake District during the same year of 1.77 mg L^{-1} (George and Charlton 1996)). Concentrations of DOC also varied and the DOC was not always coloured.

The measured $a_{TOT}(\lambda)$ for Esthwaite Water showed that, although the overall spectral structure is dominated by $a_{DYS}(\lambda)$ in the blue and $a_{H2O}(\lambda)$ in the near-IR, the green and red regions of the spectra are dominated by $a_{SES}(\lambda)$. In the absence of tripton this $a_{SES}(\lambda)$ is primarily due to absorption by phytoplankton. The combined effect of these optically active components serves to reduce the overall influence of the pigment absorption features that were visible in the $a_{SES}(\lambda)$.

The measured surface reflectance spectra showed some visible differences which could be attributed to changes in phytoplankton composition. When dominated by Cyanophytes the reflectance spectra exhibited a third peak not present in reflectance spectra dominated by other classes. The location of the main green reflectance peak also appeared to vary with dominant phytoplankton class. However, overall the reflectance spectra tended not to exhibit the strong between-class spectral differences that were present in the pure culture $R_{PHY}(\lambda)$ as measured in the experimental tank (Figure 32). This is because in the natural environment the concentrations of the various optically active water quality components can not be controlled as they could in the experimental tank. The spectral features from these components serve to mask some of the more subtle between-class differences associated with the various class distinctive marker pigments. This notwithstanding, the sixteen surface reflectance spectra and the CASI reflectance spectra from the natural water body present an opportunity to test the phytoplankton biomass and identification algorithms developed in the following chapters.

Chapter Six Chlorophyll-*a* retrieval algorithms

Chlorophyll-*a* (CHL*a*) is a frequently measured and important water quality parameter because it effectively describes the trophic status of a water body. Trophic status is a key indicator of water quality. Consequently many workers have sought to develop remote sensing algorithms to retrieve CHL*a* concentration using the strong relationship between CHL*a* concentration and reflectance (Gower *et al.* 1984, Carder *et al.* 1986, George and Hewitt 1990, Dekker *et al.* 1992a, Rundquist *et al.* 1995). In Case I waters these algorithms can be very effective because variations in the optical quality of a water body correlate strongly with variations in CHL*a* concentration. Only the package effect (§ 2.1.5.2) and variation in dominant species type detract from this relationship (Bidigare *et al.* 1989, Bricaud and Stramski 1990). In Case II waters the optical properties of the water body become more complex because of the presence of inorganic sediment and this reduces the strength of this relationship (Kirk and Tyler 1986, Quibell 1991, Schalles *et al.* 1997).

Once developed, CHL*a* retrieval algorithms can be used for the mapping of variations in phytoplankton biomass concentration across the surface of a water body. Traditionally empirical algorithms have been developed from ground truth CHL*a* measurements (Quibell 1992). These algorithms have suffered from being site and time specific, with their applicability to different water bodies or to the original water body at a different time of the year being limited. For remote sensing to become a viable option for water managers a universal algorithm is required which could be applied to all possible scenarios (Quibell 1992, Dekker 1993). Such an algorithm would have to cope with the changes in the magnitude and spectral quality of reflectance from target water bodies caused by: variations in illuminating radiance; variations in the other optically active water quality components; variations in the colour of the sediments on the lake bed (bottom reflectance); the variations in the magnitude of the reflected flux caused by large variations in CHL*a* concentration; variations in dominant and sub-dominant phytoplankton type; and variations in the growth phase of the target phytoplankton species as the intra-cellular pigments change with the oncoming of senescence as observed by Matthews and Boxall (1994).

This chapter describes the measurement of CHL*a* on water samples and the development of CHL*a* retrieval algorithms for remotely sensed reflectance spectra presented in the preceding chapters. These algorithms are developed firstly from the controlled environment of the experimental tank and then from the natural reflectance spectra measured from Esthwaite Water.

6.1 Assessment of the hot methanol extraction technique for chlorophyll-*a* analysis

The CHL*a* concentrations reported in this work have been measured using the hot methanol extraction method described by Talling and Driver (1963). After collection of the phytoplankton on a glass fibre filter (Whatman GF/C) the pigment was extracted by boiling in 10 mL Analar grade methanol for 3 minutes. The sample was centrifuged at 3000 rpm for 5 minutes and the solvent volume measured. Absorption was measured in a spectrophotometer at 750 and 665 nm. The CHL*a* concentration was calculated according to:

$$CHLa(\mu gL^{-1}) = \frac{13.9\left(OD_{(665)} - OD_{(750)}\right)v_e}{v_f P} \quad \text{.....6.1.1}$$

where $OD_{(\lambda)}$ is the optical density (absorbance) at the required wavelength, v_e is the volume in litres of extracted solvent, v_f is the volume in litres of filtered lake water and P the pathlength of the cuvette in centimetres.

The hot methanol extraction method was chosen because hot methanol is the most efficient means of extracting chlorophyll from cells (J.Talling pers. comm.). The precision of this analytical method can be assessed by analysis a series of eight multiple CHL*a* measurements made on the Esthwaite Water samples measured in the experimental tank analysis. The results of this analysis are presented in Table 13. At all measured concentrations the percentage of the mean accounted for by the standard error is relatively small, suggesting that the hot methanol extraction method can cope with a broad range of CHL*a* concentrations.

Table 13. The mean and standard error for the CHL*a* measurements made to complement the Esthwaite Water reflectance measurements made in the experimental tank.

Mean CHLa concentration (μgL^{-1})	*Number of samples*	*Standard error*	*Percentage of mean accounted for by standard error*
19.9	6	0.352	1.8%
16.2	6	0.447	2.8%
16.4	6	0.158	1.0%
127.9	6	2.052	1.6%
23.9	6	0.462	1.9%
38.1	6	0.457	1.2%
45.6	6	0.403	0.9%
25.5	6	0.331	1.3%

6.2 Development of chlorophyll-*a* retrieval algorithms from the pure culture phytoplankton reflectance spectra

The collection of pure culture laboratory-based measurements of phytoplankton reflectance presented in this study allows the development of new algorithms for the estimation of CHL*a* and for an assessment of how such algorithms are affected by variations in the dominant class.

Oceanographic remote sensing studies performed in deep water regions have tended to use reflectance ratios in the blue to green wavelengths (Morel and Prieur 1977, Andre 1992). These ratios work because the green $R_{PHY}(\lambda)$ peak increases with phytoplankton concentration through increased scattering, while the increase in scattering in the blue is mitigated by the increase in absorption by CHL*a* at these wavelengths. George and Hewitt (1990) found that a green to blue algorithm was also appropriate for the Case I lakes (§ 2.1.5) of the English Lake District. However many inland waters can be Case II where the optical influence of phytoplankton in the blue wavelengths is masked by the other optically active parameters. Consequently CHL*a* retrieval algorithms intended for inland waters tend to use the spectral effects of phytoplankton in the red wavelengths (Millie *et al.* 1992). Using linear regression Dekker *et al.* (1991) found that CHL*a* could best be predicted using a ratio of bands centred on 665 : 712 nm. Gitelson *et al.* (1993) also suggested that the far-red and near-IR regions of the spectrum are rich in information pertaining to phytoplankton biomass which can be very useful in the Case II waters. This is due to the relative absence of spectral influences from the other optically active water quality components and because of the juxtaposition of the CHL*a* absorption trough at 675 nm and a *ca.* 700 nm reflectance peak. Furthermore this near-IR reflectance peak has been shown to shift towards longer wavelengths as phytoplankton concentration increases which has formed the basis for new CHL*a* retrieval algorithms (Gitelson 1993, Matthews and Boxall 1994).

Inverting radiative transfer theory facilitates the development of analytical algorithms for the retrieval of water quality parameters. This analytical approach could potentially lead to the development of universal algorithms simultaneously applicable to Case I and Case II waters without recourse for contemporary ground truth measurements. Using spectral measurements of the inherent optical properties Dekker (1993) developed a universal algorithm for the prediction of CHL*a* which utilised the location of the far-red CHL*a* absorption peak at 676 nm and the relative absence of optically active parameters in the near-IR (706 nm) to monitor CHL*a* concentration. Millie *et al.* (1992) suggest that such an analytical algorithm would not be feasible for Case II water bodies because of the complexities of the inherent optical properties of such waters combined with the inability to accurately measure and model those properties. Consequently, along with Bukata *et al.* (1985) and Gitelson *et al.* (1993), Mille *et al.* (1992) suggest that an empirical approach to algorithm development is still appropriate for inland waters. This is particularly true if the inherent optical properties of different phytoplankton species are to be taken into account. Furthermore Bukata *et al.* (1985) and Gitelson *et al.* (1993)

suggest that measurements of *in situ* reflectance will still be necessary to calibrate analytical models which have been developed from the inherent optical properties.

The use of *in vivo* fluorescence to monitor the concentration of phytoplankton CHL*a* has been studied for thirty years (Topliss and Platt 1986, Gitelson *et al.* 1994, Babin *et al.* 1996). During the 1980's it became popular enough for sensors to be dedicated to monitoring phytoplankton biomass through fluorescence emissions with the development of the Fluorescence Line Imager (Zacharias *et al.* 1992). Fluorescence algorithms are typically based on observations at 685 nm which is where CHL*a* fluoresces.

6.2.1 Single band chlorophyll-*a* retrieval algorithms

The full set of pure culture $R_{PHY}(\lambda)$ spectra from the experimental tank work were used to assess the effectiveness of using a single wavelength observation to predict CHL*a* concentration. The strongest correlation between $R_{PHY}(\lambda)$ at a single wavelength and measured CHL*a* concentration was at $R_{PHY}(562)$ which gave a correlation coefficient of 0.463 ($R^2 = 0.214$). The correlation for the natural logarithm of both the CHL*a* and $R_{PHY}(\lambda)$ was slightly stronger at 0.713 ($R^2 = 0.508$). Such a weak relationship is unlikely to form the basis of a successful CHL*a* retrieval algorithm for multi-temporal analysis or between-lake comparisons. The relationship also suggests that the spectral response of $R_{PHY}(\lambda)$ to changes in CHL*a* concentration are quite variable for the species considered.

The strength of single band to CHL*a* relationships improve when the spectra are drawn from a single class. The strongest correlation coefficients for the classes being at: $R_{PHY}(444)$ for the Chlorophytes ($R = -0.969$, $R^2 = 0.934$); $R_{PHY}(750)$ for the Cyanophytes ($R = 0.588$, $R^2 = 0.346$); $R_{PHY}(700)$ for the Bacillariophytes ($R = 0.988$, $R^2 = 0.976$). The spectral variations in these correlation coefficients are shown in Figure 42. The relatively consistent strength of the correlation coefficients displayed by the Bacillariophytes and the Chlorophytes indicates that $R_{PHY}(\lambda)$ for the selected species respond similarly to changes in CHL*a* concentration. Wavelengths where the correlation coefficients were lower occur in regions of the spectra where the response of $R_{PHY}(\lambda)$ to increases in biomass differ for the two species measured in the experimental tank. The consistently low correlation coefficient for the Cyanophytes suggests that the *Anabaena* and *Microcystis* reflectance spectra respond differently to CHL*a* concentration as was suggested in § 4.3.5. The greater magnitude of the *Microcystis* $R_{PHY}(\lambda)$ combined with the saturated response in $R_{PHY}(\lambda)$ to increases in CHL*a* and the presence of the gas vacuoles in the *Microcystis* would appear to be the main causes of this break-down in correlation on a class level for the two Cyanophytes considered here.

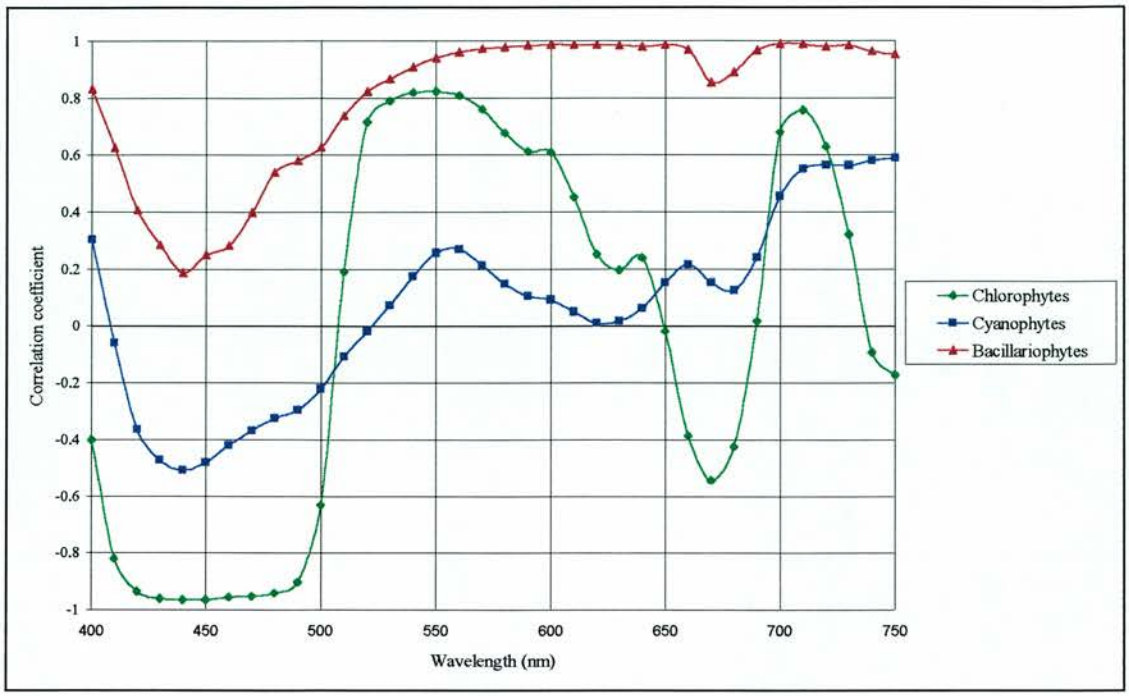

Figure 42. The correlation coefficients for $R_{PHY}(\lambda)$ for the phytoplankton classes considered in the experimental tank against measured CHL*a* concentration.

When considering spectra from a single species the correlation analysis highlights a strong relationship between $R_{PHY}(\lambda)$ and measured CHL*a* concentration at most visible wavelengths (Figure 43). From the blue wavelengths to the respective inflection points in the pure culture $R_{PHY}(\lambda)$ (as noted in the discussion of the individual species § 4.3.4) the correlation coefficients were negative. At these wavelengths increases in CHL*a* lead to reduced $R_{PHY}(\lambda)$ as the effect of increased $a_{PHY}(\lambda)$ exceeds the increase in $b_{PHY}(\lambda)$. At wavelengths longer than the respective inflection points the correlations were strongly positive as the increases in $b_{PHY}(\lambda)$ exceed the increases in $a_{PHY}(\lambda)$ resulting in a net increase in $R_{PHY}(\lambda)$. *Microcystis* shows a markedly reduced correlation coefficient at 620 nm where absorbing pigments are biliproteins not CHL*a*. This suggests that the rate of increase in absorption by PCE with biomass increase was dissimilar to the rate of increase in absorption by CHL*a*.

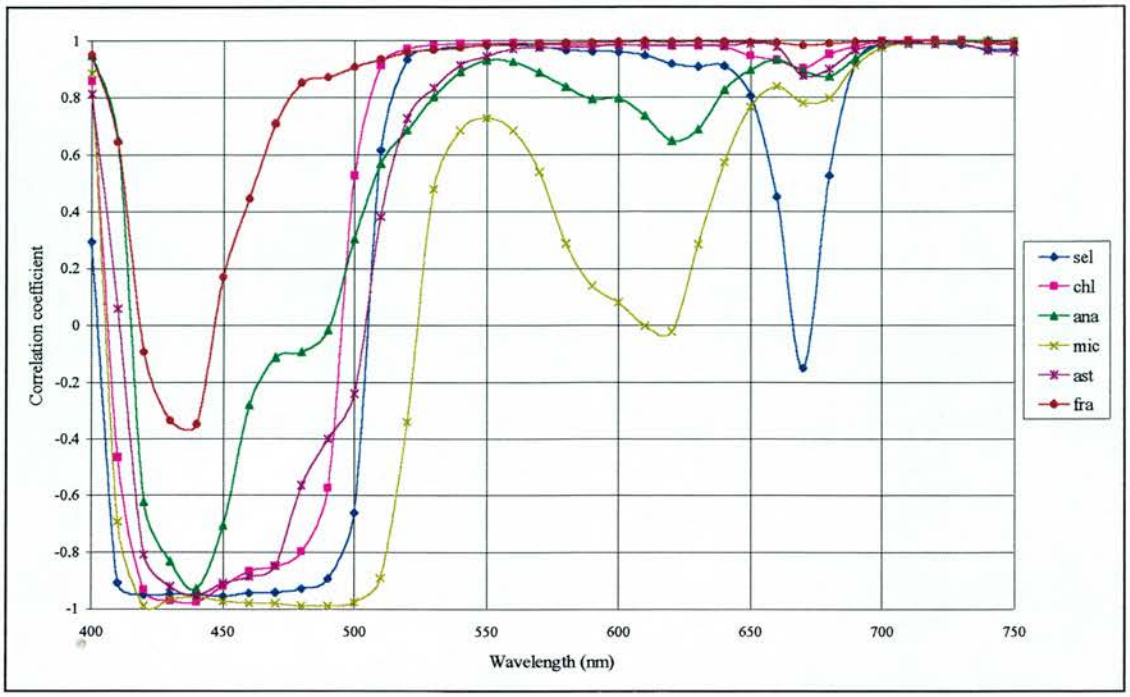

Figure 43. The correlation coefficients for $R_{PHY}(\lambda)$ for each species against measured CHL*a* concentration. For expansion of legend see Figure 20.

Figure 42 and Figure 43 show that across the visible spectrum the different phytoplankton species and classes exhibit unique relationships between $R_{PHY}(\lambda)$ and CHL*a*. As a consequence, algorithms based upon single bands, whether they be narrow or broad, may be unsuccessful in estimating CHL*a* when applied to water bodies which experience changes in species composition on a spatial or temporal basis.

6.2.2 Ratio-based chlorophyll-*a* retrieval algorithms

To determine empirically which ratio had the strongest relationship with CHL*a* concentration a Turbo Pascal v7.0 program (*Raticorr.pas*) was used (Malthus *et al.* 1995). The program is input a selection of reflectance spectra and accompanying information on a water quality parameter. It calculates the correlation coefficients for all possible ratio combinations and that parameter. The results are presented in a matrix of correlation coefficients (an R-matrix) with the ratio numerator wavelength on the *y*-axis and the ratio denominator wavelength on the *x*-axis and the correlation coefficient on the *z*-axis depicted by contrast. It is then possible to identify the reflectance ratio which best models changes in that water quality parameter, which in this instance is CHL*a* concentration.

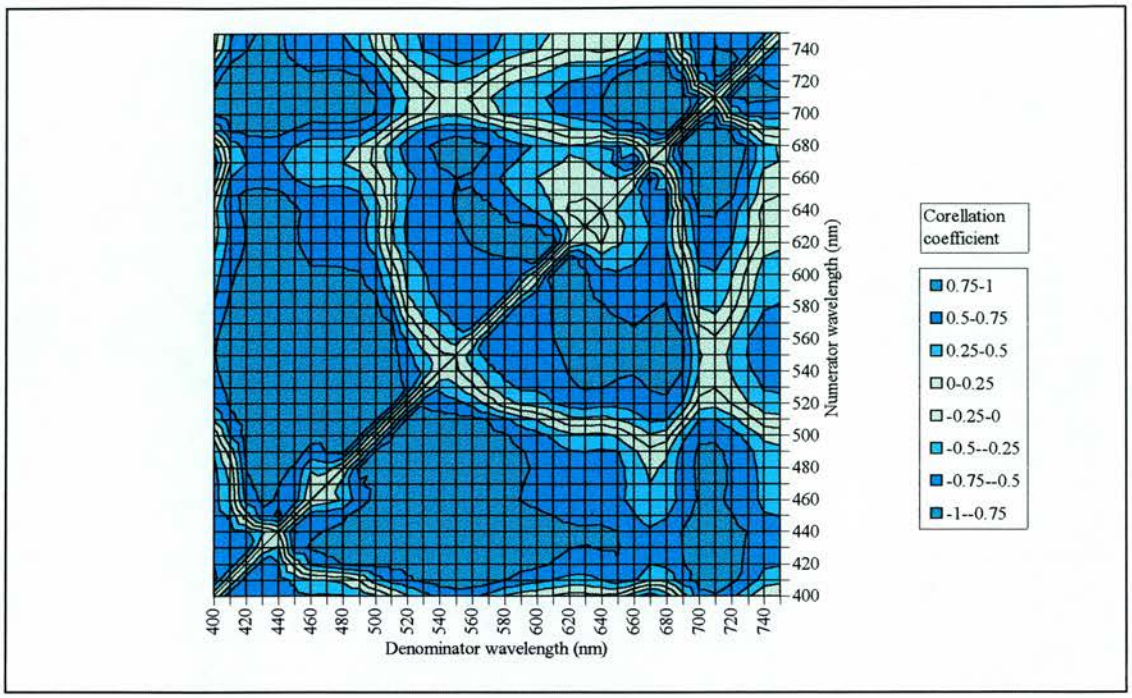

Figure 44. The R-matrix for the regression of the combined species $R_{PHY}(\lambda)$ against measured CHL*a*.

Combining all the 54 $R_{PHY}(\lambda)$ measured in the experimental tank, the program identified the ratio 702 : 678 nm is being most adept at describing changes in CHL*a* concentration (Figure 44). Regressing the reflectance ratio $R_{PHY}(702) : R_{PHY}(678)$ against CHL*a* gives:

$$\text{CHL}a\ (\mu g\ L^{-1}) = 84.9(R_{PHY}(702) : R_{PHY}(678)) - 71.49\ (R^2 = 0.961) \quad \text{.....6.2.2.1}$$

with an F-value of 1289.4 significant at the 99.9% probability level (Figure 45).

Broadening the observational bandwidth of this ratio from two to ten nm centred upon 702 and 678 nm does not reduce the predictive capabilities of Equation 6.2.2.1:

$$\text{Chl}-a\left(\mu g L^{-1}\right) = 92.36\left(R_{PHY}\left(\Sigma_{698}^{708}\right) \Big/ R_{PHY}\left(\Sigma_{682}^{672}\right)\right) - 76.89\ (R^2 = 0.959) \quad \text{.....6.2.2.2}$$

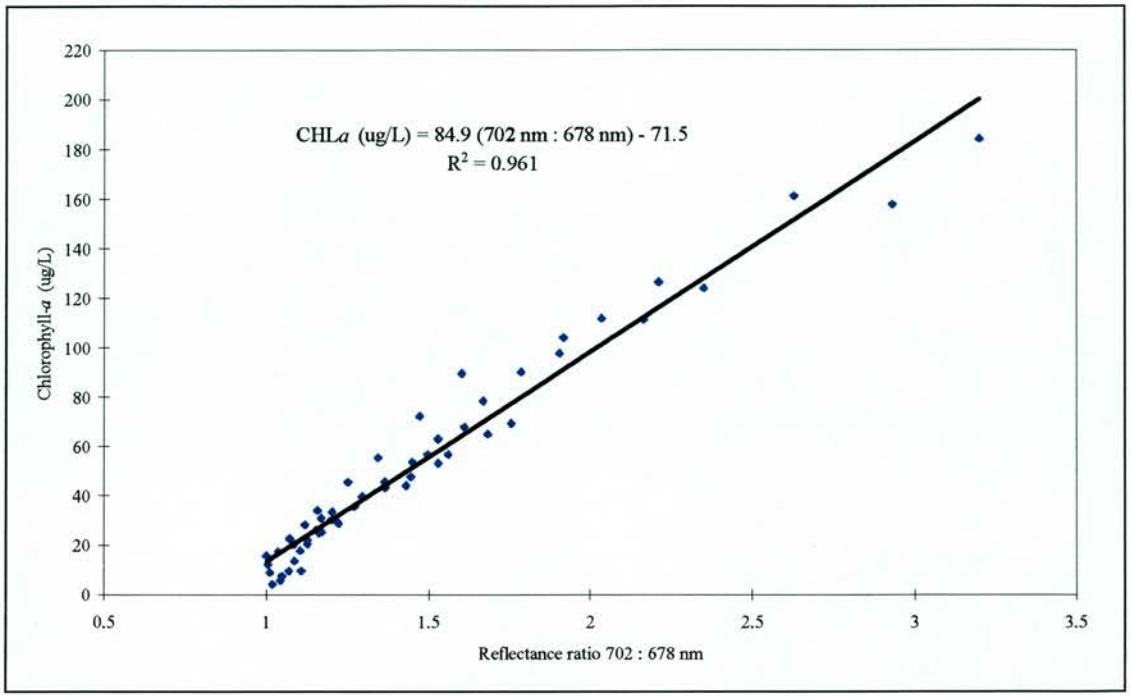

Figure 45. The linear regression for the ratio $R_{PHY}(702) : R_{PHY}(678)$ against CHL*a* (μgL^{-1}) (Equation 6.2.2.1).

Although Equation 6.2.2.1 was empirically derived it has a strong analytical basis if the components of $a_{TOT}(\lambda)$ are considered. The numerator wavelength is located at the near-IR $R_{PHY}(\lambda)$ peak. This peak has previously been attributed to CHL*a* fluorescence (Morel and Prieur 1977, Carder *et al.* 1986, Zacharias *et al.* 1992, Matthews and Boxall 1994). Phytoplankton CHL*a* fluoresces at 685 nm (Britton 1983) while the mean near-IR peak location measured in this research and reported by Gitelson *et al.* (1993) and Matthew and Boxall (1994) occurs at longer wavelengths. Furthermore phytoplankton fluoresce less than 1% of the intercepted photons and these are emitted in a random direction (Britton 1983). Hence half the emitted photons will travel in a downward direction and half in an upward direction. Only the latter will contribute to the upwelling flux. However half of these upwelling photons will be in such a trajectory as to be unlikely to penetrate the water-air interface. Thus only one quarter of the photons intercepted by phytoplankton cells will be fluoresced in such a manner as to be recorded by a remote-sensing instrument. Such a small flux is likely to be at the noise level of most radiometers so could not be successfully applied to the prediction of CHL*a* concentrations in most natural situations. Study of the inherent optical properties indicates that this peak is actually due to a minimum in $a_{TOT}(\lambda)$ at *ca.* 700 nm (Van Stokkom*et al.* 1993, Dekker 1993, Gitelson *et al.* 1993, George and Charlton 1996). This $a_{TOT}(\lambda)$ minima is transposed into the near-IR $R_{PHY}(\lambda)$ peak. In the absence of $a_{PHY}(\lambda)$ beyond 700 nm combined with a relatively constant $a_{H2O}(\lambda)$ then phytoplankton reflectance at this peak is a function of $b_{PHY}(\lambda)$ which increases with biomass, thus the magnitude of the near-IR $R_{PHY}(\lambda)$ peak is directly proportional to CHL*a* concentration.

The denominator wavelength in Equation 6.2.2.1 is located directly upon the centre of the far-red CHL*a* absorption maxima. Thus it is the site of a local $R_{PHY}(\lambda)$ trough where increased reflectance from scattering is counterbalanced by decreased reflectance through absorption, resulting in little net change in reflectance as biomass increases. The magnitude of $R_{PHY}(678)$ is therefore a function of two conflicting responses to increases in biomass: increased scattering events leading to increased reflectance; and increased absorption events leading to decreased reflectance. Thus although the magnitude of $R_{PHY}(678)$ does increase with biomass the increase is considerably less than at an absorption independent wavelength such as $R_{PHY}(702)$.

The ratio in Equation 6.2.2.1 is therefore using concomitant optical phenomena which occur as CHL*a* concentrations increase: increased reflectance at 702 nm; and decreased reflectance at 678 nm.

6.2.3 Species and class-specific ratio-based chlorophyll-*a* retrieval algorithms

Study of the residuals generated by Equation 6.2.2.1 indicates that for some species there was a degree of non-randomness of the residuals between predicted and measured CHL*a* (Figure 46).

When all the species are considered together the residuals between predicted and measured CHL*a* are random. However Figure 46 shows that on a species level there is a trend in the residuals which is typically a shift from negative to positive values as CHL*a* increases. ANOVA analysis of these residuals using first class, then species, to model the residual variation suggests that these treatments do afford some explanation. Using class as the model gave an F-value of 4.367 $p < 0.05$. Using species as the model gave an F-value of 6.784 $p < 0.01$. Hence more of the residual variation is explained by considering individual species than when considering classes. This has implications for the development of a universal CHL*a* retrieval algorithm which must be able to cope with changes in the dominant species in a target bloom over the growing season.

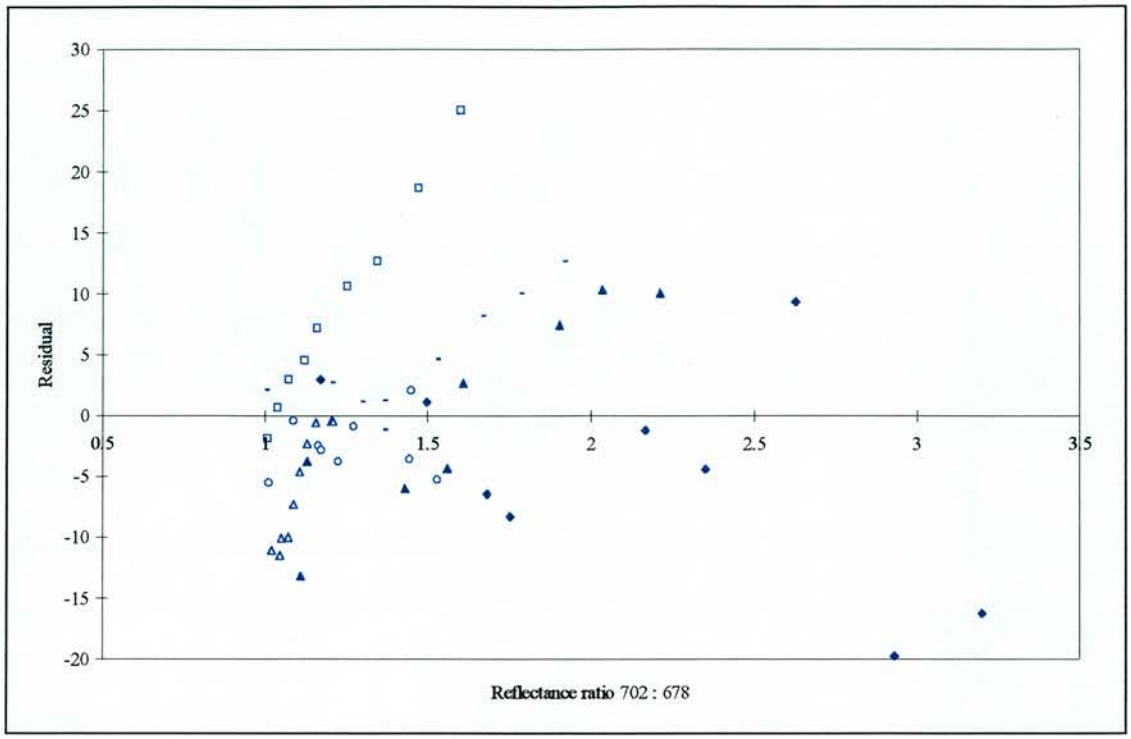

Figure 46. The residuals from the regression of the CHL*a* prediction algorithm Equation 6.2.2.1 against measured CHL*a* for the pure culture $R_{PHY}(\lambda)$. The residuals for each species are represented by different symbols: Filled diamonds = *Selenastrum*; Empty squares = *Chlorella*; Filled triangles = *Anabaena*; Short dashes = *Microcystis*; Empty circles = *Asterionella*; and Empty triangles = *Fragilaria*.

The hypothesis that a species specific algorithm based on Equation 6.2.2.1 might be more appropriate was tested by regressing the reflectance ratio $R_{PHY}(702) : R_{PHY}(678)$ against CHL*a* for the $R_{PHY}(\lambda)$ for each individual species. Table 14 shows the results of this exercise giving the slope and intercept parameters for the regression equation which maximises the R^2 value for the individual species. These parameters differ for each species and general result in an improved explanation of measured variance in CHL*a* over that offered by Equation 6.2.2.1. However Equation 6.2.2.1 explained 96% of the variance in measured CHL*a* so, although it does appear to be susceptible to error according to the dominant species type, the slight improvement in accuracy afforded by the species specific versions of Equation 6.2.2.1 are unlikely to be off-set by the increased application complexity of using such species specific algorithms.

Table 14. The regression coefficients for the reflectance ratio $R_{PHY}(702) : R_{PHY}(678)$ against CHL*a* for each species and each class.

Class or *species*	*Slope*	*Intercept*	R^2
Selenastrum	77.93	-61.27	0.979
Chlorella	127.11	-114.47	0.999
Chlorophytes	**76.23**	**-54.76**	**0.968**
Anabaena	102.21	-98.52	0.991
Microcystis	98.44	-86.67	0.993
Cyanophytes	**99.31**	**-90.94**	**0.983**
Asterionella	86.44	-75.95	0.976
Fragilaria	156.76	-156.71	0.976
Bacillariophytes	**98.16**	**-91.61**	**0.949**

Raticorr.pas was then applied to the $R_{PHY}(\lambda)$ from the individual classes and individual species to determine the most appropriate individual CHL*a* prediction algorithms. Regression analysis of the various ratios identified as having the strongest correlation coefficients are summarised in Table 15. The R-matrices describing the relationships between CHL*a* and the reflectance ratios for the individual classes all show a similar pattern to the R-matrix of the combined species (Figure 44). The R-matrix for the Chlorophytes (Figure 47) shows consistently high correlation coefficients. The R-matrix for the Cyanophytes (Figure 48) shows more ratios with low correlation coefficients especially where ratios combine $R_{PHY}(\lambda)$ observations around 660 nm with those from 450 to 630 nm. Cyanophyte $R_{PHY}(\lambda)$ ratios with high correlation coefficients when regressed against CHL*a* are those which combine wavelengths around 440 nm with any other wavelength, observations around 560 with those from 410 to 530 nm, 570 to 640 nm and those beyond 700 nm and observations around 720 nm with any other wavelength. The Bacillariophyte R-matrix (Figure 49) shows a large region of high correlation coefficients with ratios combining observations at wavelength from 400 to 550 nm, and observations from 610 to 690 nm with those from 540 to 660 nm, and ratios with observations from 700 to 750 nm with those longer than 540 nm.

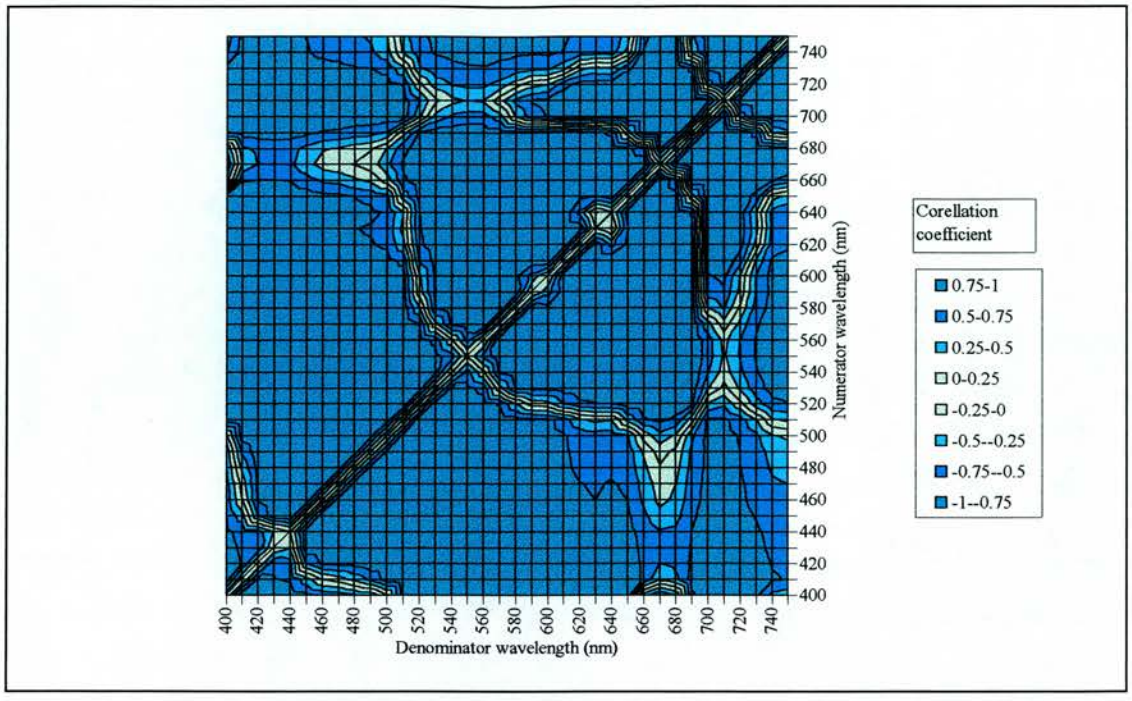

Figure 47. The R-matrix for the regression of the Chlorophyte $R_{PHY}(\lambda)$ against measured CHL*a*.

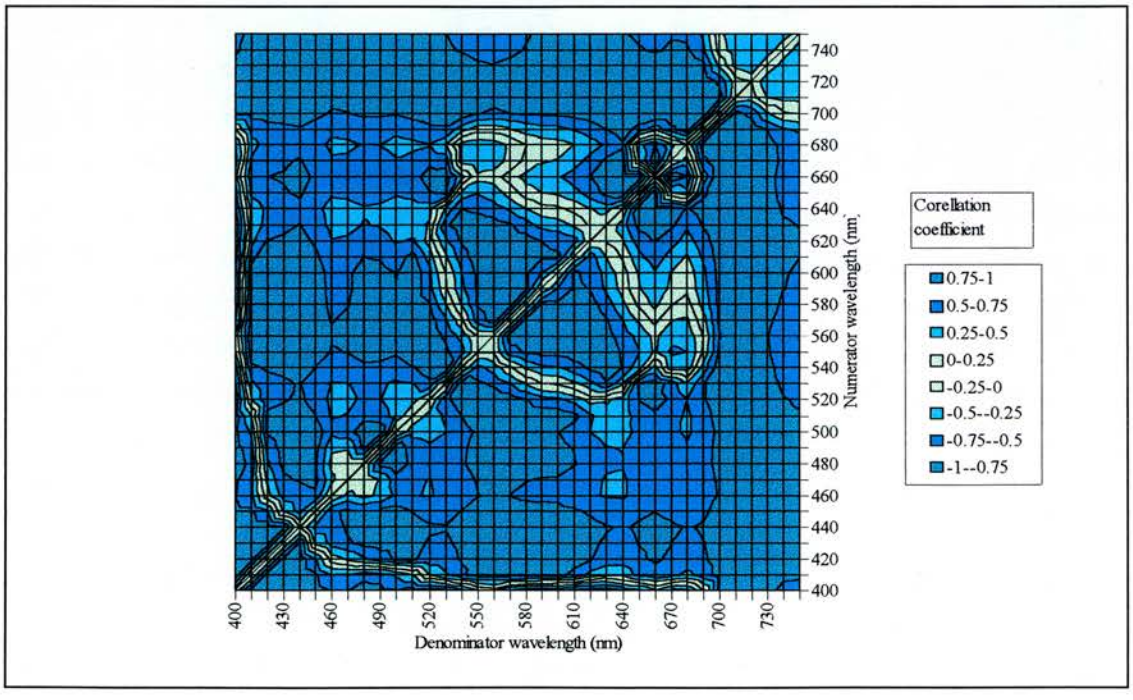

Figure 48. The R-matrix for the regression of the Cyanophyte $R_{PHY}(\lambda)$ against measured CHL*a*.

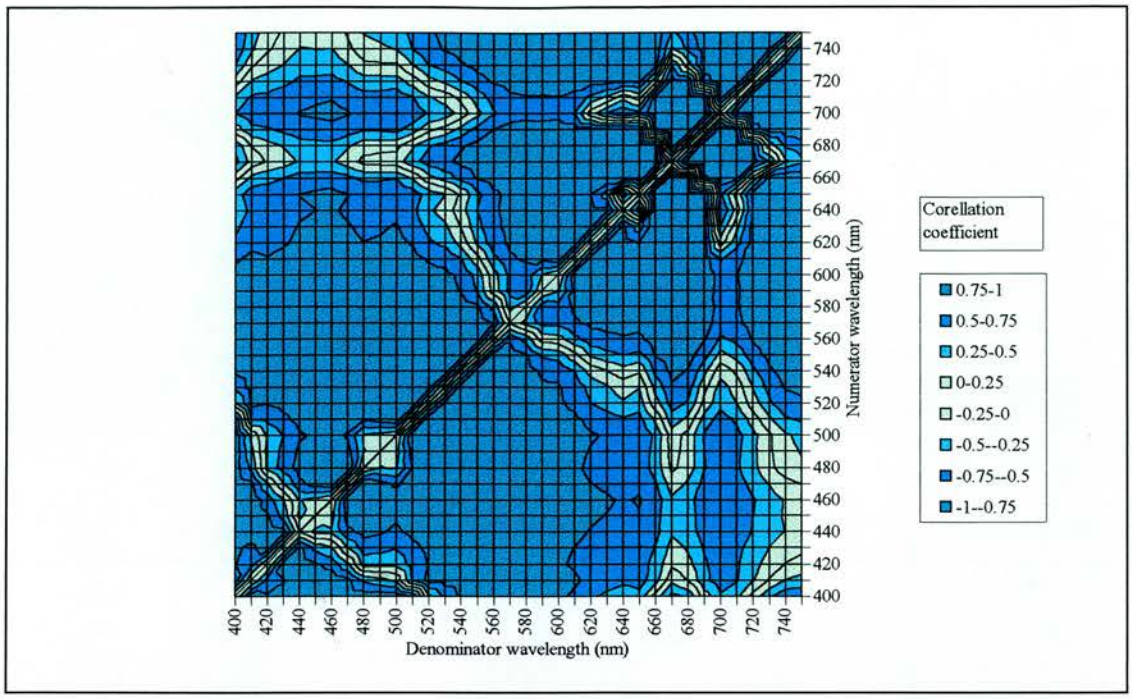

Figure 49. The R-matrix for the regression of the Bacillariophyte $R_{PHY}(\lambda)$ against measured CHL*a*.

A summary of the class specific and species-specific algorithms is presented in Table 15. All the algorithms explained a large percentage of the variance in measured CHL*a*. The algorithms for the individual species explained slightly more variance than does the 'universal' algorithm calculated using all the species spectra. The algorithms for the individual species all drew on numerator wavelengths within ± 4 nm of the 702 nm used in the universal algorithm. Similarly the species denominator wavelengths were all within ± 6 nm of 678 nm (the denominator in the universal algorithm). The denominator wavelengths were similar to the reflectance troughs identified in Table 10 which have been identified as the centre of the CHL*a* absorption band which has been shown for these particular cultures to vary slightly between species (§ 4.3.3.1). The numerator wavelengths of these algorithms correspond to the location of the near-IR $R_{PHY}(\lambda)$ peak which will be shown to be concentration dependent (§ 6.2.5). It therefore appears that, due to the reliance on the inherent optical properties of the ubiquitous phytoplankton pigment CHL*a*, these far-red : near-IR CHL*a* retrieval algorithms are not significantly altered by the species of the target bloom.

Table 15. Summary of the CHL*a* retrieval algorithms for reflectance from the individual species, classes and a combination of all species used in the experimental tank.

Target group	*Reflectance ratio*	*slope*	*intercept*	R^2	*n*
Selenastrum	704 : 682	93.05	-70.50	0.983	9
Chlorella	702 : 674	104.94	-95.93	0.999	9
Anabaena	702 : 684	109.78	-107.32	0.994	9
Microcystis	706 : 678	94.95	-76.92	0.994	9
Asterionella	696 : 676	88.83	-83.91	0.983	9
Fragilaria	698 : 672	120.09	-123.53	0.998	9
Chlorophytes	698 : 668	74.28	-63.89	0.977	18
Cyanophytes	702 : 678	99.31	-90.94	0.983	18
Bacillariophytes	690 : 674	138.60	-140.36	0.978	18
All species	702 : 678	84.90	-71.49	0.961	54

Oceanic Case I water CHL*a* retrieval algorithms typically rely upon ratios combining the blue reflectance trough (the $a_{PHY}(\lambda)$ maxima) over the green reflectance peak ($a_{PHY}(\lambda)$ minima). Inverting this ratio gives a positive correlation with CHL*a*. The R-matrix for all the $R_{PHY}(\lambda)$ spectra (Figure 44) shows that the strongest correlation coefficient for ratios combining observations of reflectance between 400 and 600 nm was for the ratio 560 : 440 nm. Linear regression of this reflectance ratio against measured CHL*a* gave:

$$\text{CHL}a\ (\mu g\ L^{-1}) = 23.07\ (R_{PHY}(560) : R_{PHY}(440)) - 14.13\ (R^2 = 0.915) \quad \text{.....6.2.2.3}$$

explaining more than 90% of the variance in measured CHL*a* (this was not improved by logging the ratio). Broadening the bandwidth to ten nm did not significantly alter the strength of this relationship:

$$\text{Chl}-a\left(\mu gL^{-1}\right) = 24.40\left(R_{PHY}\left(\Sigma_{566}^{556}\right) \Big/ R_{PHY}\left(\Sigma_{446}^{436}\right)\right) - 15.65\ (R^2 = 0.911) \quad \text{.....6.2.2.4}$$

6.2.4 Ratio-based algorithms for restricted phytoplankton biomass concentrations

Studies of Case I lakes with a range of CHL*a* concentrations have shown that different algorithms may be required for low and high CHL*a* concentration lakes, in particular the far-red : near-IR algorithm, has been found not to be applicable to oligotrophic lakes (George and Charlton 1996). To assess the requirements for low and high CHL*a* algorithms the pure culture $R_{PHY}(\lambda)$ were split into three groups < 20 μgL^{-1}, > 20 and < 60 μgL^{-1}, and > 60 μgL^{-1} which roughly correspond with oligotrophic, mesotrophic and eutrophic water bodies. *Raticorr.pas* was used to assess the most suitable ratio for each group for the prediction of CHL*a*. The results are summarised in Table 16.

Table 16. Summary of the correlation coefficients for the wavelength ratios with the strongest correlation with CHL*a* concentration when $R_{PHY}(\lambda)$ were grouped according to CHL*a* concentration.

Grouping	*Numerator (nm)*	*Denominator (nm)*	R^2	*n*
< 20 $\mu g\ L^{-1}$ CHLa	496	460	0.967	11
>20 and < 60 $\mu g\ L^{-1}$ CHLa	694	678	0.904	27
> 60 $\mu g\ L^{-1}$ CHLa	726	674	0.962	16

The ratio with the strongest propensity to predict CHL*a* for the low concentration grouping was essentially a green to blue ratio. Ratios based upon the far-red $R_{PHY}(\lambda)$ trough and the near-IR peak

did not produce strong correlation coefficients for this grouping. The medium and high concentration groupings do appear to use these $R_{PHY}(\lambda)$ features.

6.2.5 Chlorophyll-*a* retrieval algorithms based on the location of near-IR reflectance peak

Single band and ratio-based algorithms are not the only means to estimate CHL*a* concentration from $R_{PHY}(\lambda)$. The $R_{PHY}(\lambda)$ data sets show that with the exception of *Fragilaria* there is an apparent relationship between the location of the near-IR reflectance peak and the biomass concentration (see the $R_{PHY}(\lambda)$ in Figure 24 to Figure 29) which is summarised in Table 17. When individual classes are considered the location of the near-IR peak explained more than 90% of the variation in measured CHL*a* for the Chlorophytes but only 60% of the variation in measured CHL*a* for the Cyanophytes and Bacillariophytes. For all classes combined:

$$\text{CHL}a\ (\mu g\ L^{-1}) = 7.67\ (\text{near-IR}(\lambda)) - 5304.8\ (R^2 = 0.620) \quad \text{.....6.2.5.1}$$

which is shown in Figure 50. Taking the natural-log of CHL*a* did not improve this relationship. However it possible that more accurate location of the peak would improve the strength of this relationship. During the measurements of $R_{PHY}(\lambda)$ in the experimental tank the spectroradiometer was set to record at 1 nm intervals which set the level of accuracy for peak identification. A higher spectral resolution may well improve this algorithm.

Table 17. Regression of the location of the near-IR $R_{PHY}(\lambda)$ peak (nm) against measured CHL*a* (μgL^{-1}) for the individual species and classes.

Class or *species*	*Slope*	*Intercept*	R^2
Selenastrum	14.85	-10300.6	0.945
Chlorella	5.00	-3422.1	0.881
Anabaena	14.88	-10382.5	0.902
Microcystis	10.70	-7420.3	0.873
Asterionella	4.68	-3227.6	0.784
Fragilaria	7.67	-5315.0	0.346
Chlorophyte	8.50	-5852.7	0.825
Cyanophyte	10.05	-6980.0	0.709
Bacillariophyte	5.60	-3871.5	0.732
All classes	7.68	-5304.8	0.620

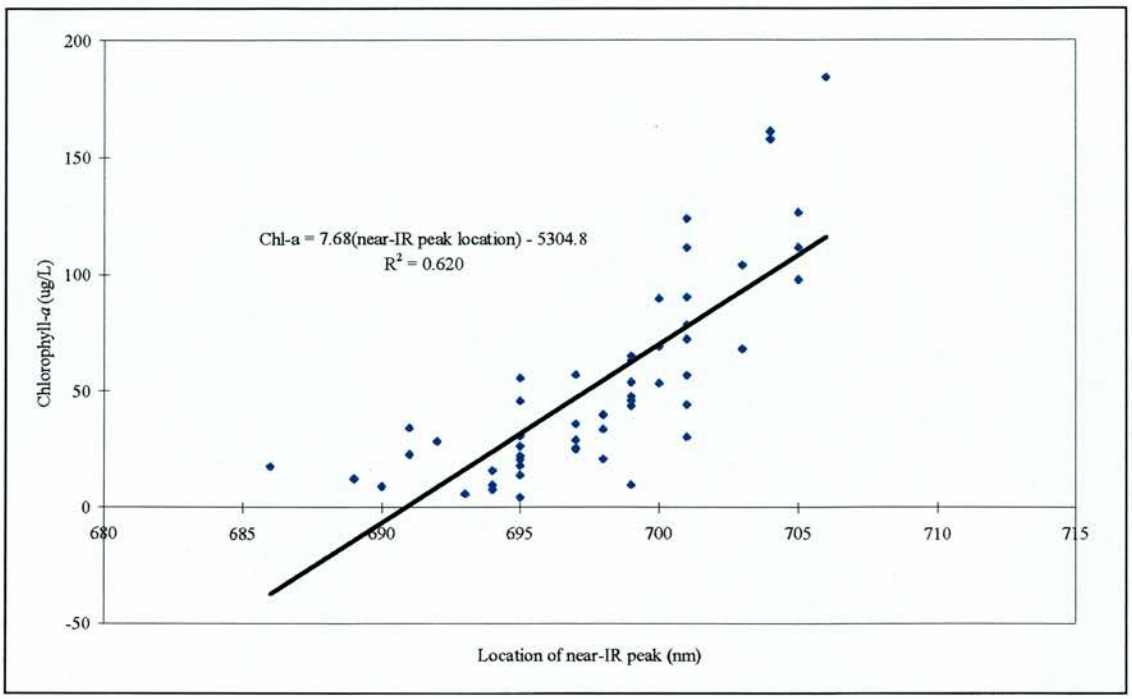

Figure 50. Linear regression of the location of the near-IR reflectance peak (nm) against measured CHL*a* concentration (μg L^{-1}).

A non-linear relationship between near-IR peak location and CHL*a* had been noted by Gitelson *et al.* (1993) and Matthews and Boxall (1994). The regression analysis performed by Rundquist *et al.* (1995) suggested a linear relationship (0.0053 near-IR peak (λ) + 1.028 with an R^2 of 0.98) for the medium and low CHL*a* concentrations, however for CHL*a* concentrations > 125 μg L^{-1} they found the relationship to be non-linear.

At the species level these linear relationships between the location of near-IR peak and the CHL*a* concentration explain 60 to 90% of the measured CHL*a* variation. Combining all classes reduces the explanation to 60% which is considerably weaker than the ratio-based algorithms. Furthermore there is considerable within-species and within-class variation in the nature of this relationship. The within-species differences in the nature of the relationship arise from the difference in the shape of the far-red CHL*a* absorption peak, specifically in the shape of the fall in $a_{PHY}(\lambda)$ from the 685 nm peak towards the longer wavelengths as it is the interaction between this and $a_{H2O}(\lambda)$ that creates the shift in the location of the $a_{TOT}(\lambda)$ minima. Figure 20 shows that the shape of this shoulder is species specific.

An analytical explanation for the relationship between peak location and CHL*a* concentration is based upon the effect of combining the $a_{PHY}(\lambda)$ and $a_{H2O}(\lambda)$ upon the near-IR $R_{PHY}(\lambda)$ peak. $a_{H2O}(\lambda)$ increases from 0.5 to more than 2.5 m^{-1} from 700 to 720 nm. The $a_{PHY}(\lambda)$, and in particular the Gaussian shaped far-red $a_{PHY}(\lambda)$ peak which is attributable solely to absorption by CHL*a*, varies in magnitude with the concentration of CHL*a*. To maintain a Gaussian form as the peak absorbance increases, the

width of the pigment absorption feature, and therefore of the far-red $a_{PHY}(\lambda)$ peak, must increase. It is this variation in peak width that is of particular importance because when it is combined with the static water absorption feature at *ca.* 700 nm it causes the local minima in the $a_{TOT}(\lambda)$ to shift towards the longer wavelengths as the CHL*a* absorption curve increases with biomass concentration. Thus the associated $R_{PHY}(\lambda)$ peak shifts towards the longer wavelengths. Figure 51 is a schematic diagram which demonstrates this relationship for the *Anabaena* $R_{PHY}(\lambda)$ by showing the effect of increasing CHL*a* on the $a_{TOT}(\lambda)$. As biomass concentration increases the location of the $a_{TOT}(\lambda)$ minima moves towards longer wavelengths. This results in a strong correlation between the location of this $a_{TOT}(\lambda)$ minima and phytoplankton biomass. As $R_{PHY}(\lambda)$ is an inverse function of $a_{TOT}(\lambda)$ this correlation transposes itself into a strong relationship between the reflectance peak location and biomass.

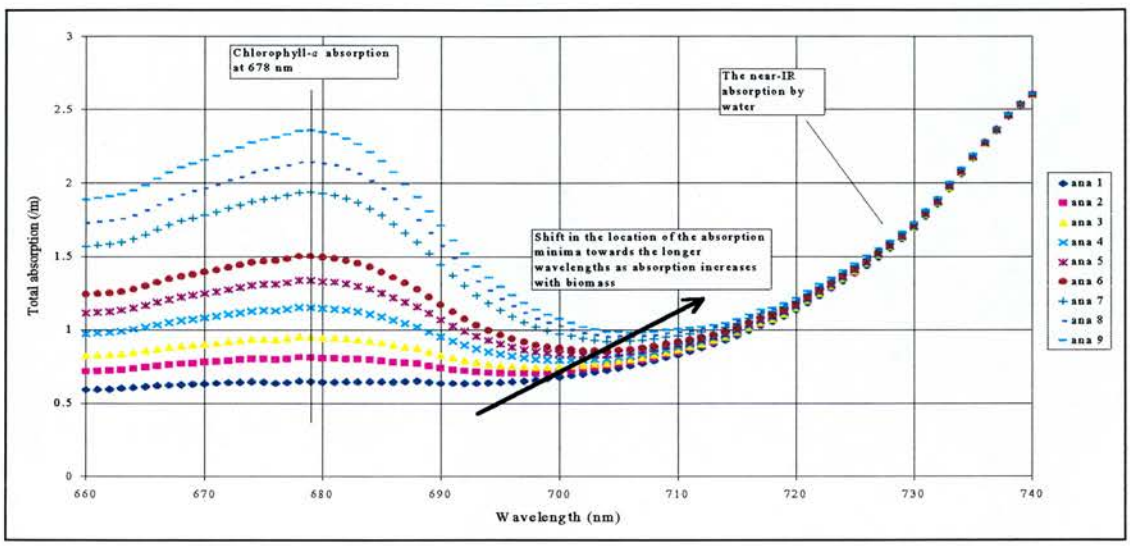

Figure 51. The shift towards longer wavelengths of the local absorption minima for the *Anabaena* $a_{TOT}(\lambda)$ positioned between the 710 nm $a_{H2O}(\lambda)$ feature and the 678 nm CHL*a* absorption peak for the biomass concentrations measured in the experimental tank.

6.3 Developing chlorophyll-*a* retrieval algorithms from the Esthwaite Water reflectance spectra

The Esthwaite Water reflectance spectra introduced in § 5.5.1 also provide an opportunity for the development of site specific CHL*a* retrieval algorithms. They are also suitable for testing the CHL*a* retrieval algorithms developed in the above section.

As was shown in § 6.2.1 single band CHL*a* retrieval algorithms tend not to be as accurate as ratio-based algorithms. As with the development of the ratio algorithms for the pure culture phytoplankton reflectance spectra, the ratio which most strongly correlated with measured CHL*a* concentration on Esthwaite Water was empirically ascertained using the *Raticorr.pas* program (§ 6.2.2). Figure 52 shows the R-matrix describing the variation in correlation coefficient by reflectance ratio. For

Esthwaite Water the spectral regions with the strongest correlation was around 700 : 670 nm. There was also an isolated region of strong correlation coefficients using reflectance ratios 570 : 590 nm.

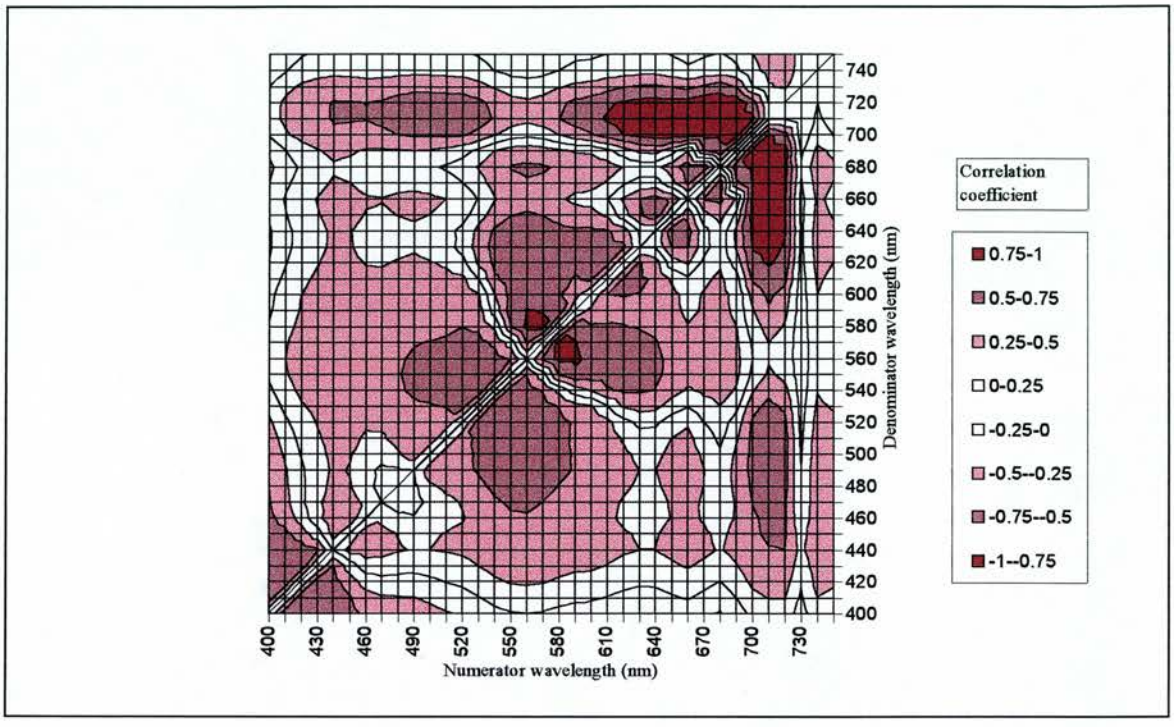

Figure 52. The R-matrix for the Esthwaite Water R(λ)(-0.01 m) against measured CHL*a* concentration.

When considering the full set of Esthwaite Water R(λ)(-0.01 m) the strongest relationship identified was based on the near-IR reflectance peak and the far-red reflectance trough (Figure 53):

$$CHLa\ (\mu gL^{-1}) = 103.5\ (R(722) : R(684)) - 62.86\ (R^2 = 0.908) \quad \text{.....6.3.1}$$

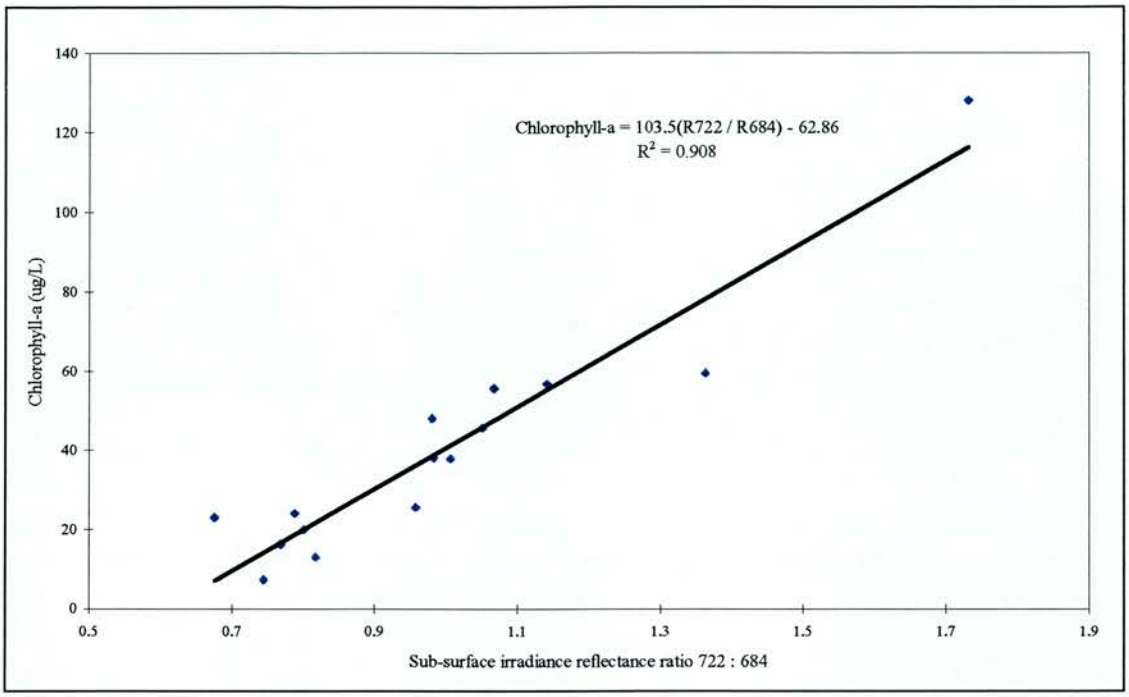

Figure 53. Regression analysis of the reflectance ratio 722 : 684 against measured CHL*a* ($\mu g\ L^{-1}$) for the Esthwaite Water reflectance spectra.

As was the case for Equation 6.2.2.1 this ratio varies with CHL*a* because the concomitant changes in $a_{PHY}(\lambda)$ and $b_{PHY}(\lambda)$ which occur as CHL*a* concentrations increase. The combination of increased absorption and increased scattering means that any ratio where the numerator is drawn from reflectance between 670 and 690 nm, and the denominator from reflectance between 700 and 730 nm will correlate positively with CHL*a* concentration in Case I type waters. This is useful when considering the broad-band airborne imagers as opposed to the hyperspectral radiometers used in this study. Ten nanometer broad-bands centred on 680 nm and 720 nm give the following strong relationship:

$$Chl-a(\mu gL^{-1}) = 89.86\left(\Sigma_{675}^{685} \Big/ \Sigma_{715}^{725}\right) - 55.51 \quad (R^2 = 0.867) \quad \text{.....6.3.2}$$

Analysis was undertaken to assess whether residuals between CHL*a* predicted by Equation 6.3.1 and measured CHL*a* for the ground truth samples could be explained by the variation in any of the other measured water quality parameters. The residuals are shown in Table 18. The correlation between the residuals and total cell numbers was 0.508. None of the other water quality parameters were strongly correlated with the residuals (the correlation coefficient for the residuals and DOC was -0.028; and for the residuals and g_{440} was -0.318; and for the residuals and dry weight was 0.051; and for residuals and Secchi disk depth was -0.329). Notably the correlation coefficient between residuals and species diversity index was low (-0.158).

Table 18. The residuals (in ascending order) between the measured CHL*a* and predicted CHL*a* by Equation 6.3.1. The species diversity index, the dominant phytoplankton species and class at the time of sampling as identified in the surface water samples are also given.

Residual	*Date*	*Total cell numbers*	*Phytoplankton diversity*	*Dominant species*	*Dominant class*
-18.86	14-Jul-94	11236	0.75	*Oscillatoria*	Cyanophyte
-10.75	10-Oct-95	602	1.56	*Anabaena*	Cyanophyte
-8.87	10-Apr-96	3710	0.85	*Asterionella*	Bacillariophyte
-6.99	23-May-95	3196	0.50	*Sphaerocystis*	Chlorophyte
-3.51	21-Sep-94	11521	0.96	*Anabaena*	Cyanophyte
-0.73	15-Aug-95	21259	1.10	*Oscillatoria*	Cyanophyte
-0.57	20-Jun-95	4506	1.43	*Rhodomonas*	Cryptophyte
-0.30	12-Sep-95	11630	0.96	*Oscillatoria*	Cyanophyte
-0.26	04-Jul-95	7862	1.11	*Anabaena*	Cyanophyte
-0.18	06-Jun-95	1963	0.53	*Cryptomonas*	Cryptophyte
1.44	06-Jul-94	6230	0.69	*Anabaena*	Cyanophyte
5.21	01-Aug-95	15822	1.14	*Anabaena*	Cyanophyte
7.97	28-Jul-94	16299	0.31	*Anabaena*	Cyanophyte
9.38	01-Sep-94	12804	1.16	*Anabaena*	Cyanophyte
11.62	25-Jul-95	105604	0.61	*Anabaena*	Cyanophyte

It can be seen from Table 18 that Equation 6.3.1 tended to under predict CHL*a* when the reflectance spectra was dominated by *Anabaena*. This is the only indication that a dominant species type is having an influence on the accuracy of Equation 6.3.1 as a CHL*a* prediction algorithm.

Using the complete Esthwaite Water reflectance data set the strongest regression analysis relating near-IR peak location to CHL*a* explained 60% of the variance in measured CHL*a*. Removing the outlying CHL*a* sample for 25/07/95 improves this relationship to :

$$\text{CHL}a\ (\mu g L^{-1}) = 2.96 \text{ near-IR peak}(\lambda) - 2042.7\ (R^2 = 0.875) \quad \text{6.3.3}$$

The linear regression of this relationship is shown in Figure 54. Equation 6.3.3 explains considerably more of the variance in measured CHL*a* than that offered by the near-IR peak shift CHL*a* retrieval algorithm developed from the pure culture $R_{PHY}(\lambda)$ measured in the artificially controlled conditions of the experimental tank (Equation 6.2.5.1).

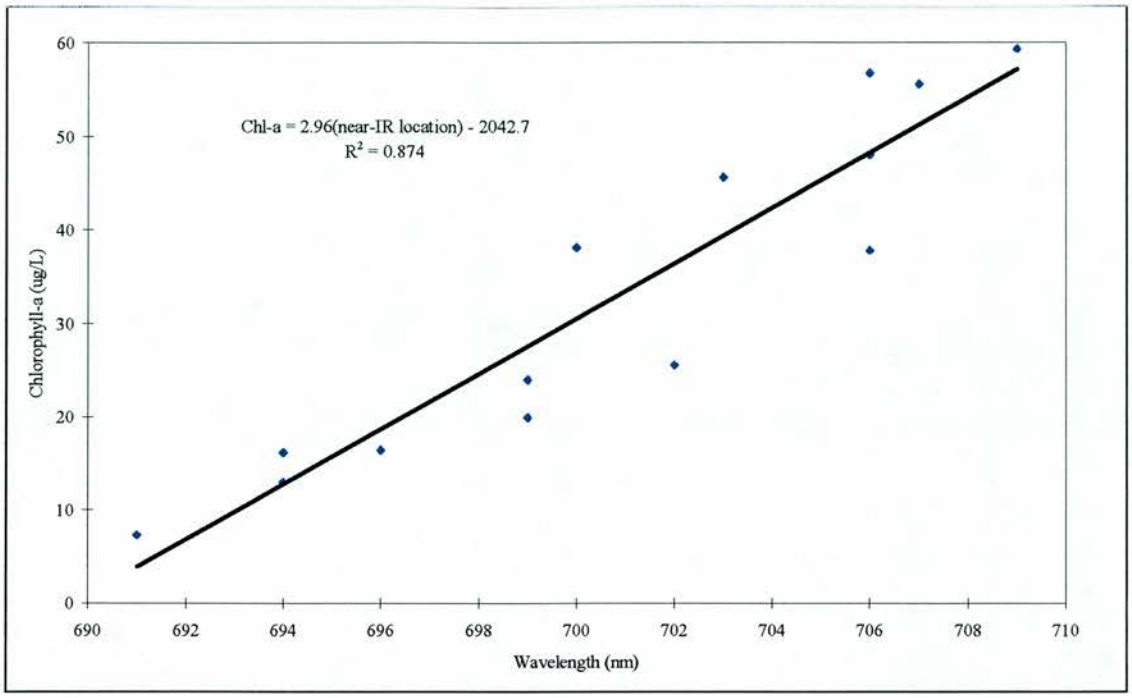

Figure 54. The regression of the location of the near-IR reflectance peak location against CHL*a* concentration. Note the outlying reflectance spectra for 25/07/94 has been excluded from this regression analysis.

6.4 Comparison of chlorophyll-*a* retrieval algorithms applied to the Esthwaite Water reflectance spectra

Table 19. Summary of the accuracy of the CHL*a* retrieval algorithms when applied to the sixteen Esthwaite Water sub-surface irradiance reflectance spectra.

Reference	*CHLa (μgL^{-1}) algorithm*	R^2	*Comments*
Ratio algorithms			
Equation 6.3.1	103.5 (722 nm : 684 nm) - 62.9	0.91	Esthwaite Water ratio
Class specific Equation 6.2.2.1	see Table 15	0.65	class tailored experimental tank ratio
Equation 6.2.2.1	84.9 (702 nm : 678 nm) - 71.5	0.63	best experimental tank ratio
Equation 6.2.2.3	23.1 (560 nm : 440 nm) - 14.1	0.16	green to blue ratio
Universal algorithms			
George and Charlton (1996)	$98.2\left(\Sigma_{705}^{695}:\Sigma_{670}^{660}\right)-71.9$	0.71	
Dekker (1993)	66.5 (706 nm : 676 nm) - 48.2	0.65	
Gitelson (1993)	a (700 nm : 675 nm)^b	0.59	see text for symbols
Peak shift algorithms			
Equation 6.3.3	2.96 near-IR peak (λ) - 2042.7	0.87	Esthwaite Water peak shift
Equation 6.2.5.1	7.67 near-IR peak (λ) - 5304.8	0.87	experimental tank peak shift

Equation 6.3.1 (empirically developed from the Esthwaite Water R(λ)(-0.01 m)) was the most successful algorithm for predicting CHL*a* concentrations in Esthwaite Water. When applied to the same data set the universal CHL*a* retrieval algorithm of Dekker (1993) (CHL*a* (μgL^{-1}) = 66.5 (R(706)

/ R(676)) - 48.2) explained 64% of the variance in measured CHL*a*. The universal algorithm of Gitelson (1993), which provides an adjustment for the trophic status of the water body (CHL*a* (μgL^{-1}) = $a(R_{700} / R_{675})^b$), explained 58% of the measured variation in CHL*a* (*a* and *b* are the adjusting factors utilising previously measured CHL*a* concentrations *a* = 1.46 (mean chlorophyll) +1.68 and *b* = -0.013 (mean chlorophyll) + 3.5). George and Charlton (1996) used an empirically derived ratio-based algorithm to map CHL*a* distribution with ten nanometer bands centred on 700 : 665 nm. When applied to Esthwaite Water data set this empirically derived ratio explained 71% of the variance in measured CHL*a*.

Table 19 shows that of the reported universal algorithms, the one with the strongest propensity to predict CHL*a*, explaining 71% of the variance in measured CHL*a*, was that of George and Charlton (1996). This algorithm was developed using reflectance spectra from lakes in the English Lake District which included Esthwaite Water. The universal algorithms of Dekker (1993) and Gitelson (1993) only explained 65% and 59% of the variance in measured CHL*a* respectively.

When the algorithm developed from the combined pure culture $R_{PHY}(\lambda)$ (Equation 6.2.2.1) was applied to the Esthwaite Water R(λ)(-0.01m) it explained 63% of the variance in measured CHL*a*, compared with an 91% explanation for Equation 6.3.1. Applying the parameters which optimised Equation 6.2.2.1 for each of the three classes studied in the experimental tank (Table 14) did not improve the accuracy of the CHL*a* prediction. Similarly, applying the class specific ratio-based algorithms (Table 15) only improved the explanation of variance by 2% to 65%.

When far-red to near-IR algorithms developed for the experimental tank (Equation 6.2.2.1) and Esthwaite Water (Equation 6.3.1) were used to predict the CHL*a* concentration from the sixteen Esthwaite Water reflectance spectra, the lake specific algorithm explained 91% of the variance while the experimental tank algorithm only explained 63%. Even when the experimental tank algorithm was tailored to the dominant class present in Esthwaite Water at the time of sampling it could only explain 65% of the variance in CHL*a* concentration.

The near-IR peak shift algorithms developed from both the experimental tank reflectance spectra (Equation 6.2.5.1) and the Esthwaite Water reflectance spectra (Equation 6.3.3) were less precise at predicting CHL*a* concentration than was Equation 6.3.1. However they performed better than the other ratio-based algorithms presented in Table 19. Interestingly they both had R^2 values of 0.87 hence explained a similar level of variance in measured CHL*a*.

When the strongest ratio-based CHL*a* retrieval algorithm (Equation 6.3.1) was applied to the Esthwaite Water reflectance spectra acquired with the CASI at an altitude of 1000 m flying in hyperspectral mode (§ 5.5.2) it explained 83% of the variance in the ground truth measurements of

CHL*a* concentration. The peak shift CHL*a* retrieval algorithm (Equation 6.3.3) only explained 46% of the variance.

6.5 Discussion of the chlorophyll-a retrieval algorithms

The single band reflectance algorithms developed in § 6.2.1 are subject to the different response in reflectance that arises from the individual phytoplankton species as biomass increases. This response is controlled by their inherent optical properties which are a function of their cellular morphology and intracellular pigment content. When empirically calculated using reflectance spectra from the same species, the precision of single band algorithms at predicting phytoplankton concentration was good. However their applicability as universal algorithms, to be applied to natural water bodies where phytoplankton mixtures may be present, is limited because the most appropriate wavelength for retrieving CHL*a* concentration differs for each class (Figure 42).

The reflectance ratio algorithms presented in this study have been developed using the R-matrix technique. The R-matrix simultaneously highlights those ratios with the potential to predicted CHL*a* concentration and those ratios which are unsuccessful as predictors of CHL*a*. As such the R-matrix is a particularly useful tool for the development of ratio reflectance algorithms.

The success of Equation 6.3.1 for predicting CHL*a* concentration from both surface and aerially acquired Esthwaite Water reflectance spectra can be attributed to the fact that this algorithm was derived from the surface water reflectance spectra for Esthwaite Water. It has therefore been optimised for the particular optical conditions encountered. With the exception of the George and Charlton (1996) algorithm, the numerator wavelength (722 nm) for Equation 6.3.1 is both longer and further displaced from the numerator wavelength than for the universal ratio algorithms considered in Table 19 (700 to 706 nm). The George and Charlton (1996) algorithm has an equivalent displacement between the numerator and the denominator but has a numerator observation at 700 nm. Consideration of the $a_{TOT}(\lambda)$ for the Esthwaite Water samples (Figure 39) shows the variance in $a_{TOT}(\lambda)$ to be greater closer to wavelengths around 700 nm. At 722 nm this variance was minimal. In the absence of $a_{PHY}(\lambda)$ at this denominator wavelength the direct relationship between $b_{PHY}(\lambda)$ and $R_{PHY}(\lambda)$ as phytoplankton biomass increases (§ 6.2.2) is not interfered with by the concurrent indirect relationship between $a_{PHY}(\lambda)$ and $R_{PHY}(\lambda)$. This is not the case for the numerator wavelength where $a_{PHY}(\lambda)$ is present.

The ratio-based algorithm developed from the Esthwaite Water reflectance spectra (Equation 6.3.1) was developed from the same water body over two growing seasons. It has therefore been developed from reflectance spectra during high and low CHL*a* concentrations and when different species were in ascendance. It therefore has some multi-temporal capabilities, being able to deal with a range of biomass concentrations and a range of different phytoplankton species.

The algorithm developed from the pure culture reflectance spectra (Equation 6.2.2.1) has some inherent universal properties in that it is applicable to the reflectance spectra from the three phytoplankton classes (and six phytoplankton species) used in the experimental tank. It was shown that the accuracy of Equation 6.2.2.1 can be improved if the algorithm parameters are tailored to accommodate the appropriate class or species. However the benefits of this improved accuracy are low, whilst the increased complexity of assessing a reflectance spectra for dominant phytoplankton type and applying the appropriate algorithm are is likely to preclude the application of these class specific algorithms. When applied to a natural water body (as opposed to the artificial environment of the experimental tank) the accuracy of Equation 6.2.2.1 is reduced.

Although tailoring Equation 6.2.2.1 to phytoplankton class or species type did lead to small increases in accuracy over the generic CHL*a* prediction algorithm developed for the pure culture $R_{PHY}(\lambda)$, these small increases in accuracy were not realised when these species specific algorithms were applied to the Esthwaite Water reflectance spectra. It would appear that the presence of a variety of species and the problems of performing reflectance measurements in the natural environment decreases the effectiveness of these species specific algorithms.

Algorithms have also been presented based on the location of the near-IR reflectance peak. These algorithms were developed from both the laboratory and from the lake reflectance spectra. Given the success of the peak shift algorithm developed from the experimental tank spectra at predicting the Esthwaite Water CHL*a* concentrations, it would appear that this form of algorithm has potential as a universal algorithm. Furthermore, application of algorithms based upon locating reflectance peaks requires hyperspectral data which is not always available from airborne remote sensing instruments. Another restriction to the application of the peak shift algorithms, which is not apparent from the data presented in this study, is the lack of an identifiable near-IR peak in the reflectance spectra from oligotrophic waters (George and Charlton 1996). Also, when being flown in hyperspectral mode, the bandwidth of the CASI is approximately 7 nm which may not be sufficient to accurately identify the location of the near-IR peak.

6.6 Conclusions for the Chlorophyll-*a* retrieval algorithms

Ratio algorithms are more successful than single band algorithms for predicting CHL*a* concentration in natural waters. The empirically derived CHL*a* retrieval algorithm Equation 6.3.1 was the most successful at predicting the CHL*a* concentration from the Esthwaite Water reflectance spectra. To assess its universal applicability this algorithm requires testing on the reflectance spectra from a range of water bodies, including the more complex spectral signatures from Case II waters.

The gain in accuracy that could be achieved by tailoring CHL*a* retrieval algorithms to the dominant species in the target bloom in natural waters is minimal and would be outweighed by the advantages of being able to apply a single CHL*a* retrieval algorithm to all fresh water bodies. The relatively poor performance of the class tailored algorithms suggests that further development of both the species specific algorithms and the class specific algorithms is required if there are to be significant improvements in prediction accuracy over site specific, empirically derived algorithms.

In that the empirical algorithm developed on Esthwaite Water reflectance spectra outperformed the universal algorithms presented by Dekker (1993) and Gitelson (1993), it would appear that, for a specific natural water body, the universal algorithms were not as successful as empirically derived CHL*a* retrieval algorithms. Consequently, if accurate CHL*a* measurements are to be achieved using remote sensing, there will always be a need for ground truth measurements from which these empirical algorithms can be developed. The R-matrix method would appear to be a particularly succinct method of establishing the most appropriate empirical algorithm from a set of ground truth reflectance spectra.

Chapter Seven Phytoplankton class identification routines

In this chapter four routines for the identification of phytoplankton class from reflectance spectra are presented. In the first section a simple ratio-based routine is presented which uses observations at two wavelengths to determine the relative abundance of classes. In the second section a canonical discriminant analysis routine is employed for the purposes of identifying the phytoplankton class from spectra. The canonical discriminant analysis routine uses a handful of observations for classification. In the third section derivative analysis is used for the identification of pigment features in spectra, the presence or absence of which can then be used to identify phytoplankton class. Calculation of derivatives requires hyperspectral data. The final section builds on the derivative analysis using a neural network to perform the classification.

All these routines are based on the same principal: identification of class marker pigment absorption features in either $a_{PHY}(\lambda)$ or in $R_{PHY}(\lambda)$ leading to the identification of the phytoplankton class. Pigment absorption features are identified by referencing against those pigment absorption properties reported in the literature and collated in the pigment database. All the routines are developed on the pure culture $a_{PHY}(\lambda)$ introduced in the previous chapter and tested on the mixed-culture $a_{PHY}(\lambda)$, the techniques are then transposed to the pure culture $R_{PHY}(\lambda)$. This procedure is theoretically sound because absorption is an inherent optical property and in the absence of significant scattering there is a strong inverse relationship between $a_{PHY}(\lambda)$ and $R_{PHY}(\lambda)$, and, because of the low volumes of pure culture required for the $a_{PHY}(\lambda)$ measurements compared to the volumes required in the experimental tank for the $R_{PHY}(\lambda)$ measurements, it is less labour intensive. However the main benefit of using an inherent optical property to develop identification routines as opposed to using reflectance is that the resulting routines should be more widely applicable to natural water bodies under various illumination conditions because they will have a pseudo-analytical as opposed to a purely empirical basis.

7.1 Phytoplankton class identification ratios

Identification of phytoplankton on the basis of *in vivo* $a_{PHY}(\lambda)$ depends upon the pigment composition of the cells and on how many wavelengths are measured (Johnsen *et al.* 1994). French *et al.* (1967) suggest that there are a comparatively small number of unique and definitive class marker pigments each with unique absorption characteristics. It should therefore be possible to successfully identify phytoplankton from observations of absorption or reflectance at a handful of wavelengths located to detect the presence of these marker pigments. Most current airborne sensors only have the ability to measure spatially in a limited number of bandsets, hence there is a need to develop identification techniques using low spectral input data for the purpose of mapping phytoplankton distributions

(Curran 1994). Reducing the phytoplankton class identification to a simple ratio would therefore be more useful to water managers with access to the current set of airborne sensors. This section assesses the potential of simple ratios, based on wavelengths located at the absorption centres of known marker pigments, to identify phytoplankton species type.

7.1.1 Considerations for a ratio-based identification routine

The advantage of using ratios over single bands for algorithm development are twofold. Firstly, in the absence of a trending $b_{PHY}(\lambda)$ spectra, the effects of phytoplankton concentration upon the optical measurement will be experienced at both wavelengths and will therefore be negated when a ratio is calculated; were a single wavelength used, the concentration effect would have to be accommodated. Secondly, any instrument or data-processing error present at all wavelengths would not affect a ratio-based routine; a single wavelength observations would be susceptible to such errors. However using multi-band ratios does increase the disrupting effect of signal noise. For this reason the number of spectral bands used in an identification algorithm should be kept to a minimum (Dekker 1993, Dekker and Peters 1993).

A successful ratio-based class identification routine would be applicable to any of the following measurements of phytoplankton optical properties: absorption, reflectance and upwelling radiance. The latter is important because it is not always possible to convert remote sensing data to reflectance (George and Charlton 1996). Ratios are not dependent upon having accurate absolute measurements of the optical property requiring only that the spectral quality has been accurately recorded.

There are some regions of the visible spectrum that should be avoided because they are the absorption centres for atmospheric components which can confuse the spectral signature (Gregg and Carder 1990). Furthermore, the shorter visible wavelengths should be avoided because shorter wavelengths are subjected to higher atmospheric interference due to Rayleigh scattering (Monteith and Unsworth 1990).

7.1.2 Methods for formulating class identification ratios

The class identification ratios discussed below were developed from pure culture $a_{PHY}(\lambda)$ the measurement of which is discussed in § 4.2.3. 88 $a_{PHY}(\lambda)$ from 19 different phytoplankton species covering five different classes (34 Chlorophyceae, 23 Cyanophyceae, 17 Bacillariophyceae, 8 Cryptophyceae and 6 Euglenophyceae) were reduced to 155 observations at 2 nm intervals from 400 to 700 nm. The ability of absorption ratios to discriminate between combinations of two classes was assessed by performing an analysis of variance (ANOVA). ANOVA uses F-ratios to describe the degree of explanation afforded by the predictor variable (class) for the predicted variable (the absorption ratio). The F-ratio is the ratio of predicted to non-predicted variance (Hope 1968).

A Turbo Pascal (v7) program (*Fratio.pas*) was written to perform an ANOVA, according to Porkess (1988), to explain the observed variation at every possible absorption ratio from 400 to 700 nm using an index of class as the predictor variable. High F-ratios indicate that the absorption ratio was suitable for between-class discrimination. The F-ratio values were output from the programme in a matrix from which the ratio with the highest F-ratio was identified. Such F-ratio matrices are similar to the R-matrices introduced in § 6.2.2. The numerator wavelength is shown on the *y*-axis and the denominator wavelength on the *x*-axis and the strength of the F-ratio on the *z*-axis depicted by contrast. The F-ratio matrix was a useful starting point for identifying ratios suitable for classifying spectral measurements.

7.1.3 Results of the phytoplankton absorption identification ratios

In the first instance the $a_{PHY}(\lambda)$ from combinations of two classes were input into *Fratio.pas*. The F-ratio matrices for combinations of Chlorophytes, Cyanophytes and Bacillariophytes (the three main classes used in this study) are presented in Figure 55, Figure 56 and Figure 57. The results of all two class combinations are summarised in Table 20.

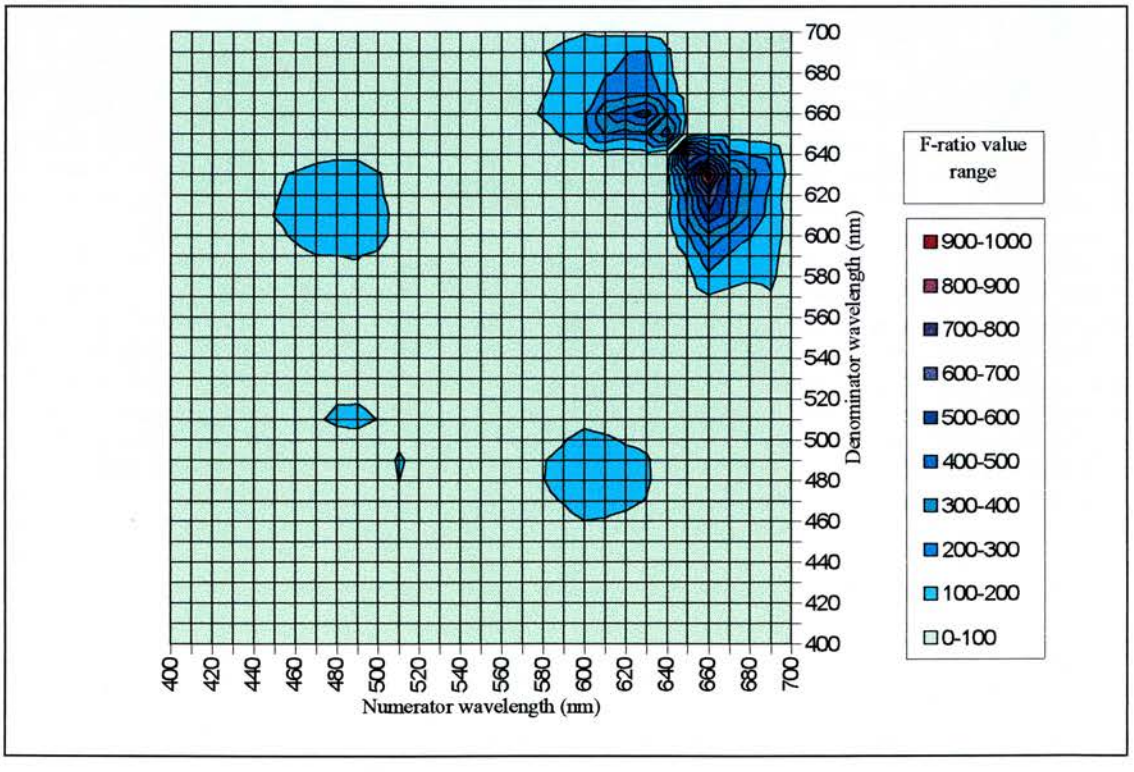

Figure 55. The Chlorophyte and Cyanophyte $a_{PHY}(\lambda)$ F-ratio matrix.

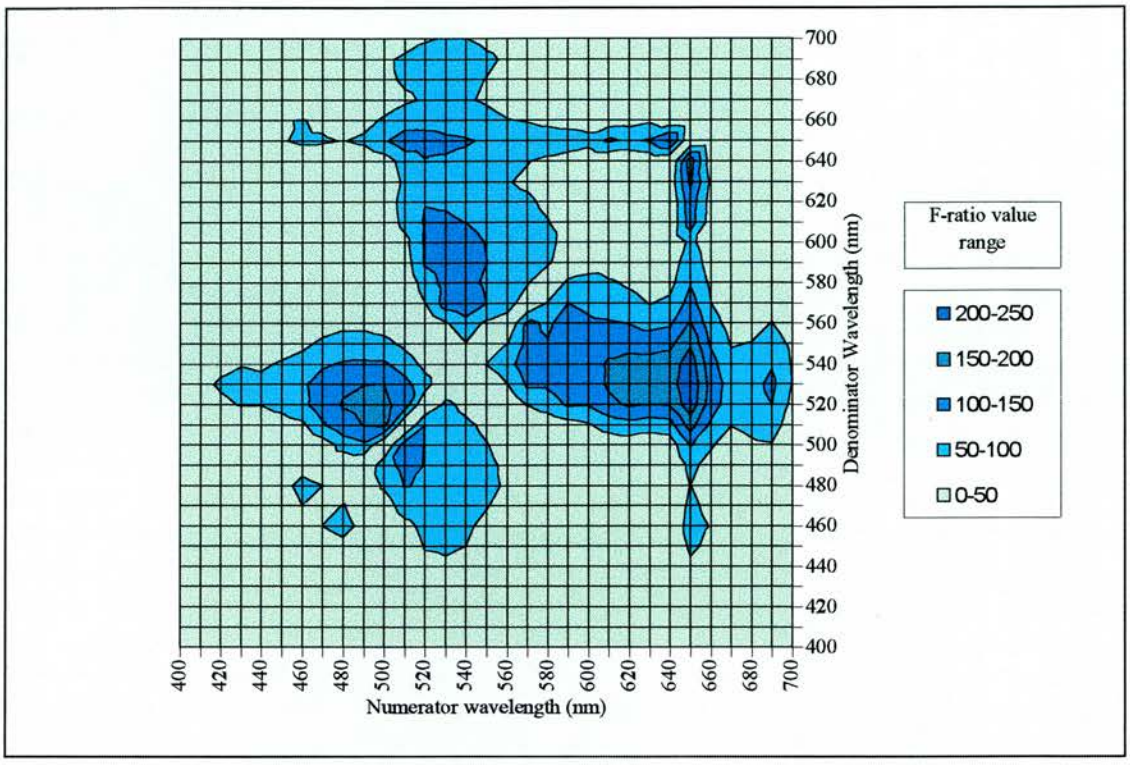

Figure 56. The Chlorophyte and Bacillariophyte $a_{PHY}(\lambda)$ F-ratio matrix.

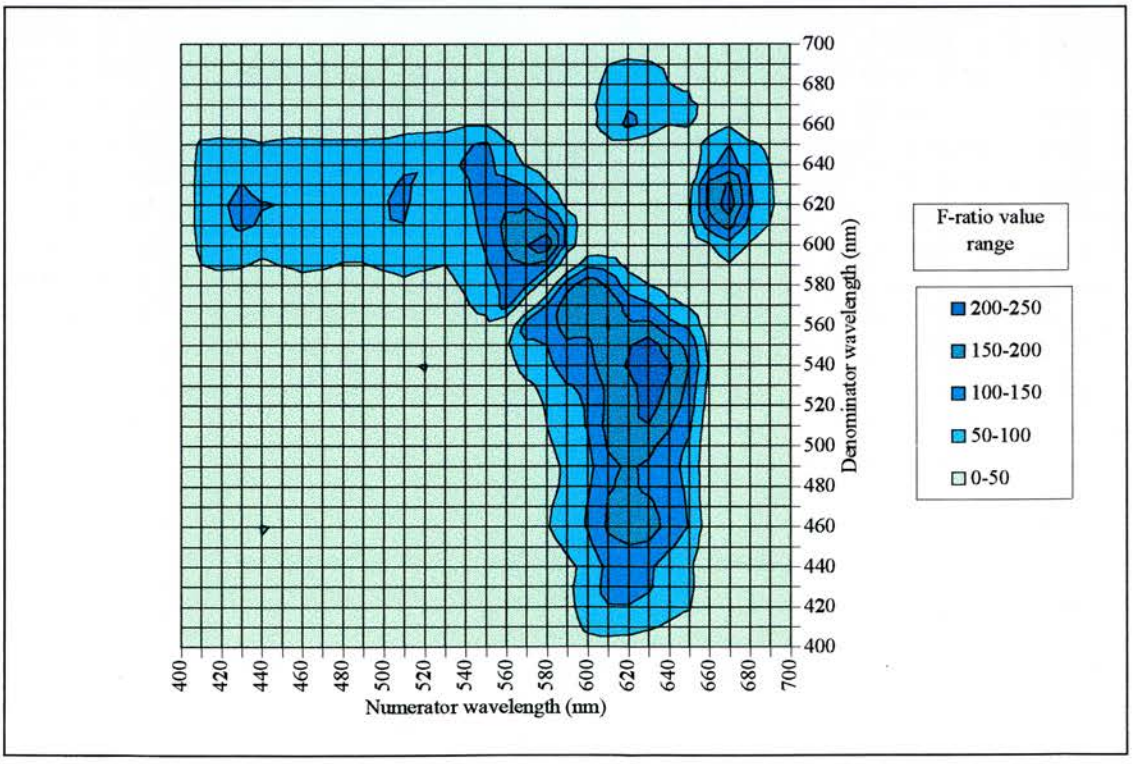

Figure 57. The Cyanophyte and Bacillariophyte $a_{PHY}(\lambda)$ F-ratio matrix.

Table 20. The $a_{PHY}(\lambda)$ ratios which best discriminate between two classes, with their F-ratios value from the ANOVA test and the number of $a_{PHY}(\lambda)$ for each class. All F-ratios significantly exceed their respective critical F-values at > 99.9% probability level.

Class 1	*Class 2*	*Numerator wavelength*	*Denominator wavelength*	*F-ratio value*	*n Class 1*	*n Class 2*
Chlorophyte	Cyanophyte	634	656	1066.6	34	23
Chlorophyte	Bacillariophyte	528	648	250.3	34	17
Cyanophyte	Bacillariophyte	542	628	244.5	23	17
Chlorophyte	Cryptophyte	572	658	495.8	34	8
Chlorophyte	Euglenophyte	492	484	11.9	34	6
Cyanophyte	Cryptophyte	580	610	504.6	23	8
Cyanophyte	Euglenophyte	658	634	805.2	23	6
Bacillariophyte	Cryptophyte	550	516	94.5	17	8
Bacillariophyte	Euglenophyte	630	530	217.9	17	6
Cryptophyte	Euglenophyte	506	490	318.5	8	6

The ANOVA analysis for the Chlorophyte and Cyanophyte classes (Figure 55) gave the strongest overall F-ratio of 1066.6 located in a compact region of high F-ratios confined to a spread of wavelengths from 580 to 700 nm. The Chlorophyte and Bacillariophyte grouping (Figure 56) gave a lower maximum F-ratio but the spread of high F-ratio values was broader with absorption coefficients from 530 to 650 nm discriminating between class when ratioed against most of the other wavelengths in the visible spectrum. The Cyanophyte and Bacillariophyte grouping gave an F-ratio matrix (Figure 57) which highlighted ratios combining wavelengths between 600 and 650 nm against those from 400 to 590 nm.

The ability of the $a_{PHY}(\lambda)$ ratios identified in Table 20 to separate two class combinations of Chlorophyceae, Cyanophyceae and Bacillariophyceae are graphically represented using box plots in Figure 58, Figure 59 and Figure 60. Reckhow (1980) used box plots to convey information graphically regarding data spread and variation. The box plots simultaneously indicate the median (the middle horizontal), interquartile range (the upper and lower horizontals), standard deviation (the extent of the notches), data range lying outside the interquartile range (the whiskers extending above and below the upper and lower horizontals), and the outlying data points (the * and 0 markers). The box plot for the Chlorophyte and Cyanophyte $a_{PHY}(\lambda)$ shows that, with a few marked exceptions, the a_{PHY} 634 : 656 nm ratio was adroit at separating the two classes (Figure 58). For both classes this ratio gave a narrow data spread and with no overlap between the groups. The box plot for the Chlorophytes and Bacillariophytes showed that the a_{PHY} 528 : 648 nm ratio tightly groups the majority of the Chlorophyte $a_{PHY}(\lambda)$ while but is less affective at grouping the Bacillariophyte $a_{PHY}(\lambda)$ (Figure 59). The a_{PHY} 524 : 628 ratio is also displayed as a box plot showing the good separation of the Cyanophytes and Bacillariophytes (Figure 60).

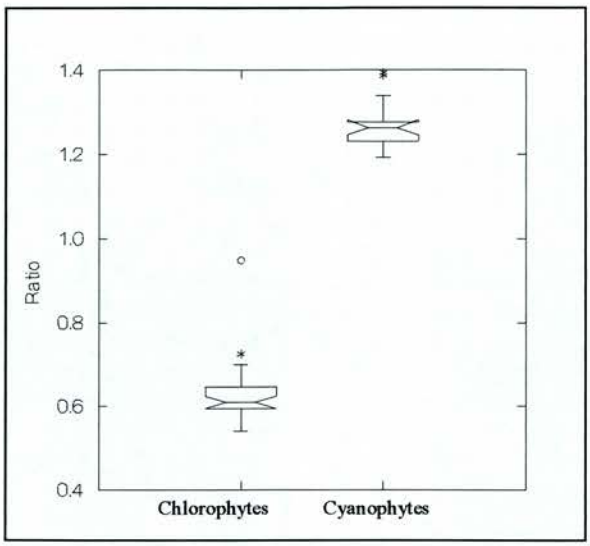

Figure 58. Box plot of the a_{PHY} ratio 634 : 656 nm for the discrimination of the Chlorophytes and Cyanophytes. See text for explanation of box plots features.

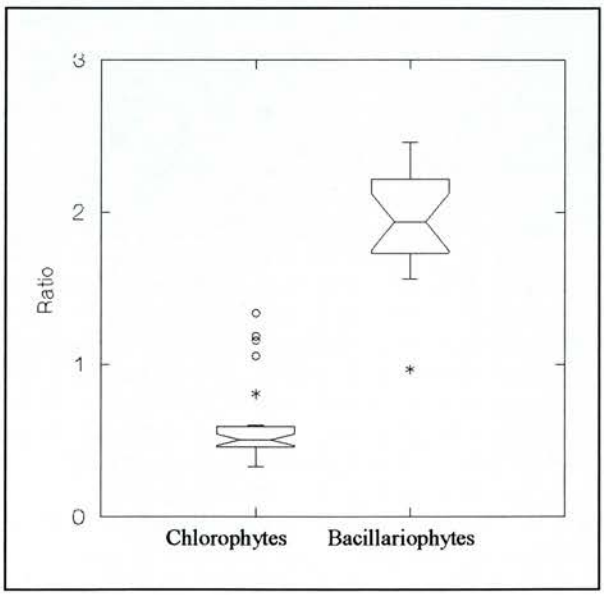

Figure 59. Box plot of the a_{PHY} ratio 528 : 648 nm for the discrimination of Chlorophytes and Bacillariophytes. See text for explanation of box plot features.

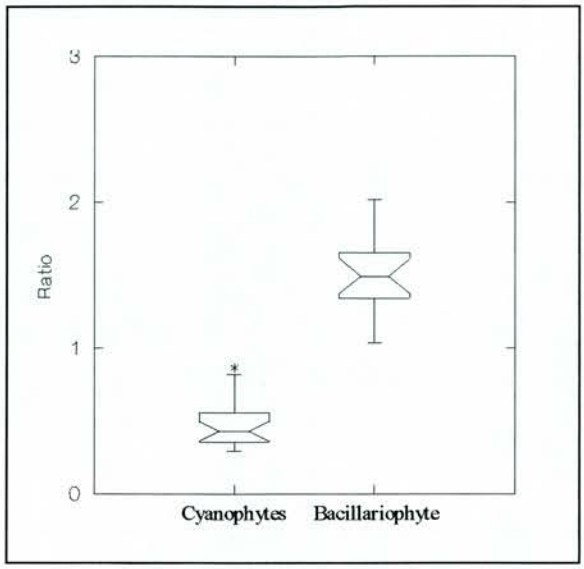

Figure 60. Box plot of the a_{PHY} ratio 542 : 628 nm for the discrimination of Cyanophytes and Bacillariophytes. See text for explanation of box plot features.

7.1.4 Discussion of the phytoplankton absorption identification ratios

The $a_{PHY}(\lambda)$ ratios listed in Table 20 were successful in discriminating between the phytoplankton groups because they utilised differences in absorption at wavelengths generated by the presence or absence of specific pigments. Pigments with absorption maxima most closely matching the wavelengths of the empirically derived $a_{PHY}(\lambda)$ ratios were identified using the pigment database and are given in Table 21. The Cyanophyte and Chlorophyte a_{PHY} ratio uses 634 and 656 nm, which are the absorption centres for PCC and CHL*b* respectively. PCC is present in Cyanophytes while absent from Chlorophytes; conversely CHL*b* is present in Chlorophytes while absent from Cyanophytes. The $a_{PHY}(\lambda)$ ratios with strong F-ratios are therefore using the absorption centres for two mutually exclusive pigments for the two classes concerned, as shown in Table 21. In the case of the Euglenophyte and Cryptophyte absorption ratio the only pigment reported as absorbing at the Euglenophyte wavelength is PCC. However the presence of biliproteins has not previously been reported in the Euglenophyceae.

Table 21. The $a_{PHY}(\lambda)$ ratios which most strongly discriminated between combinations of two class from the five classes used in this study. Class marker pigments responsible for these discriminations are also given.

Class 1	*Class 2*	*$a_{PHY}(\lambda)$ ratio*	*Class 1 marker pigment absorbing at the numerator wavelength*	*Class 2 marker pigment absorbing at the denominator wavelength*
Cyanophyceae	Chlorophyceae	634 : 656	PCC	CHL*b*
Bacillariophyceae	Chlorophyceae	528 : 648	FCX	CHL*b*
Cryptophyceae	Chlorophyceae	572 : 658	PCE	CHL*b*
Euglenophyceae	Chlorophyceae	492 : 484	DDX	CHL*b*
Cyanophyceae	Bacillariophyceae	542 : 628	FCX	PCC
Cyanophyceae	Cryptophyceae	610 : 580	PCE	PCE
Euglenophyceae	Cyanophyceae	658 : 634	CHL*b*	PCC
Cryptophyceae	Bacillariophyceae	550 : 516	PCC	FCX
Euglenophyceae	Cryptophyceae	630 : 530	PCC?	FCX
Cryptophyceae	Euglenophyceae	506 : 490	PCE	DDX

Cursory examination of Table 21 shows that 18 different absorption observations would be necessary to determine phytoplankton class on the basis of $a_{PHY}(\lambda)$ ratios. Some of the $a_{PHY}(\lambda)$ ratios are using observations of the same pigments but at slightly different locations. For example, ratios using the pigment CHL*b* to distinguish between Chlorophytes and other classes use three different wavelengths all centred around 654 nm. These slight variations in optimal wavelength are a result of the empirical nature by which these absorption ratios were established. Having ascertained which pigments are responsible for the class separations, it is possible to rationalise these wavelengths to reduce the number of observations required making the technique more applicable to those airborne sensors which have a limited number of available bands.

These rationalised wavelengths are proposed on the basis of the marker pigments *in vivo* absorption properties as recorded in the pigment database and are given in Table 22. Using these rationalised absorption wavelengths does reduce the strength of the ANOVA. The Cyanophyte and Chlorophyte F-ratio value is reduced to 966.3, the Chlorophyte and Bacillariophyte F-ratio value is reduced to 187.2 and the Cyanophyte and Bacillariophyte F-ratio value is reduced to 220.2, all of which significantly exceed their critical F-values. However, the box plots for these rationalised two class $a_{PHY}(\lambda)$ ratios (not presented) indicate that they still achieve good within-class clustering and between-class separation.

Table 22. The marker pigments and proposed observational wavelengths for the five phytoplankton classes used in this study.

Phytoplankton class	*Marker pigment*	*Observation wavelength*
Euglenophyceae	DDX	490
Bacillariophyceae	FCX	530
Cryptophyceae	PCE	576
Cyanophyceae	PCC	630
Chlorophyceae	CHL*b*	656

The ratios reported above use narrow band inputs. Airborne sensors like the Airborne Thematic Mapper (ATM) sample radiance in broader wavebands (Dekker and Peters 1993). Broadening the bands of either the $a_{PHY}(\lambda)$ or the $R_{PHY}(\lambda)$ ratios to 10 nm did not significantly reduce their propensity to discriminate between-classes.

7.1.4.1 Transposing the absorption ratios to reflectance ratios

Having identified the $a_{PHY}(\lambda)$ ratios with the strongest propensity to discriminate between two different classes and shown that those ratios rely upon wavelengths where class marker pigments are reported to absorb, it is important to ascertain whether the ratios can be transposed to $R_{PHY}(\lambda)$, thereby ensuring that they have potential as a remote sensing tool. For this purpose the pure culture $R_{PHY}(\lambda)$ introduced in § 4.3.4 were used. Note that direct transferral of the $a_{PHY}(\lambda)$ ratios to $R_{PHY}(\lambda)$ relies on the first approximation assumption that $R_{PHY}(\lambda)$ is the inverse of $a_{PHY}(\lambda)$ and that scattering by phytoplankton has little or no wavelength dependency.

The $R_{PHY}(\lambda)$ from the experimental tank were restricted to three different class types therefore only $R_{PHY}(\lambda)$ ratios dealing with Chlorophytes, Cyanophytes and Bacillariophytes were considered. The $a_{PHY}(\lambda)$ ratio wavelengths were applied to the $R_{PHY}(\lambda)$ data set and the results are summarised in Table 23. The F-ratio values for the two class $R_{PHY}(\lambda)$ ratios involving the Chlorophytes are all significant at > 99.9% level of probability but are less so than for the $a_{PHY}(\lambda)$ ratios (Table 20), while the Cyanophyte and Bacillariophyte two class $R_{PHY}(\lambda)$ ratio gives an F-ratio value greater than for the $a_{PHY}(\lambda)$ ratio.

Table 23. The two class $R_{PHY}(\lambda)$ ratios using the wavelengths identified from the $a_{PHY}(\lambda)$ ratios (Table 20) and the $R_{PHY}(\lambda)$ ratios with the highest F-ratios. The F-ratios all significantly exceed their critical F-values at the 99.9% level of probability.

Phytoplankton classes	*Best $a_{PHY}(\lambda)$ ratio as an $R_{PHY}(\lambda)$ ratio*	*F-value*	*n*	*Best $R_{PHY}(\lambda)$ ratio*	*F-value*	*n*
Chlorophyte : Cyanophyte	656 : 634 nm	196.7	18	678 : 666 nm	317.8	18
Chlorophyte : Bacillariophyte	648 : 528 nm	183.4	18	590 : 588 nm	286.7	18
Cyanophyte : Bacillariophyte	628 : 542 nm	463.1	18	588 : 600 nm	609.4	18

To ascertain whether the strongest $a_{PHY}(\lambda)$ ratio coincided with the strongest $R_{PHY}(\lambda)$ ratio *Fratio.pas* was applied to combinations of the two classes. The F-ratio matrices from this analysis are shown in Figure 61, Figure 62 and Figure 63. The strongest F-ratio for the two class $R_{PHY}(\lambda)$ ratios did not match the $a_{PHY}(\lambda)$ ratios in terms of numerators and denominators (Table 23). The Chlorophyte and Cyanophyte $R_{PHY}(\lambda)$ ratio with the highest F-ratio value was 678 : 666 nm (Figure 61). However there was a region of $R_{PHY}(\lambda)$ ratios with strong between-class discriminatory properties centred around 630 : 650 nm which was highlighted in Figure 55 as being $a_{PHY}(\lambda)$ ratios with strong discrimination properties. The strongest F-ratio for the Chlorophyte and Bacillariophyte $R_{PHY}(\lambda)$ ratio was 590 : 588

nm (Figure 62). Again there was a localised maxima within the region centred around 530 : 650 nm which was identified by the $a_{PHY}(\lambda)$ F-ratio matrix (Figure 56) as being highly discriminatory for these two classes. The strongest F-ratio for the Cyanophyte and Bacillariophyte $R_{PHY}(\lambda)$ ratio was at 588 : 600 nm. There was also a local maxima at 530 : 630 nm (Figure 63) which was the region of the highest F-values identified in the $a_{PHY}(\lambda)$ F-ratio matrix for these two classes (Figure 57).

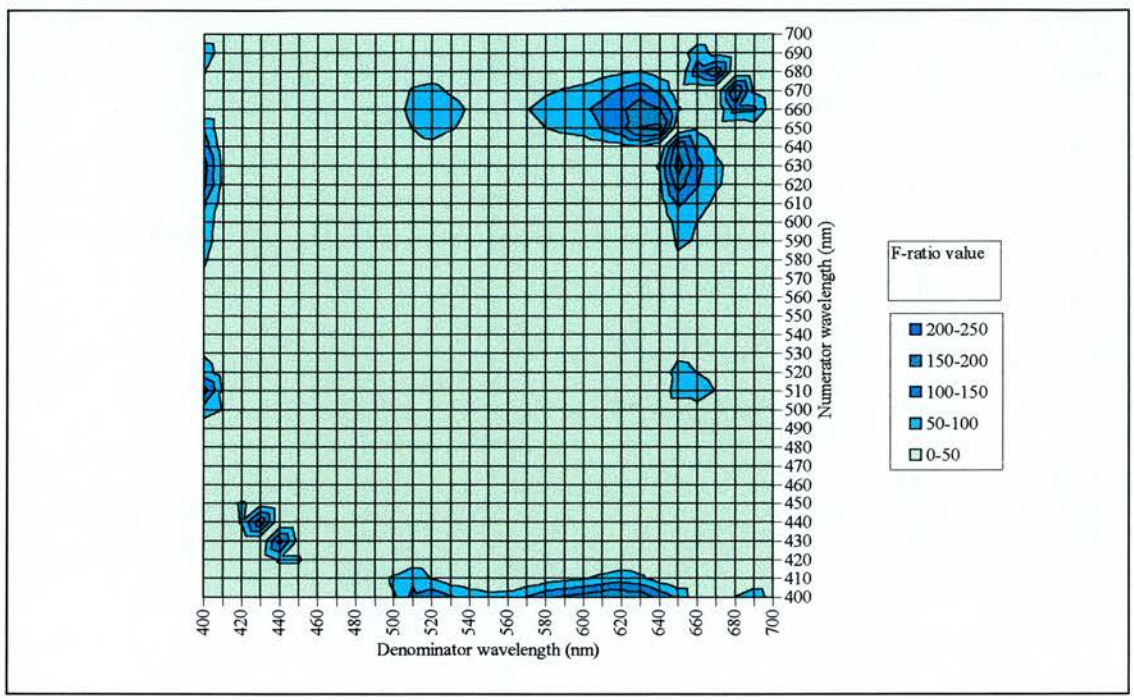

Figure 61. The Chlorophyte and Cyanophyte $R_{PHY}(\lambda)$ F-ratio matrix.

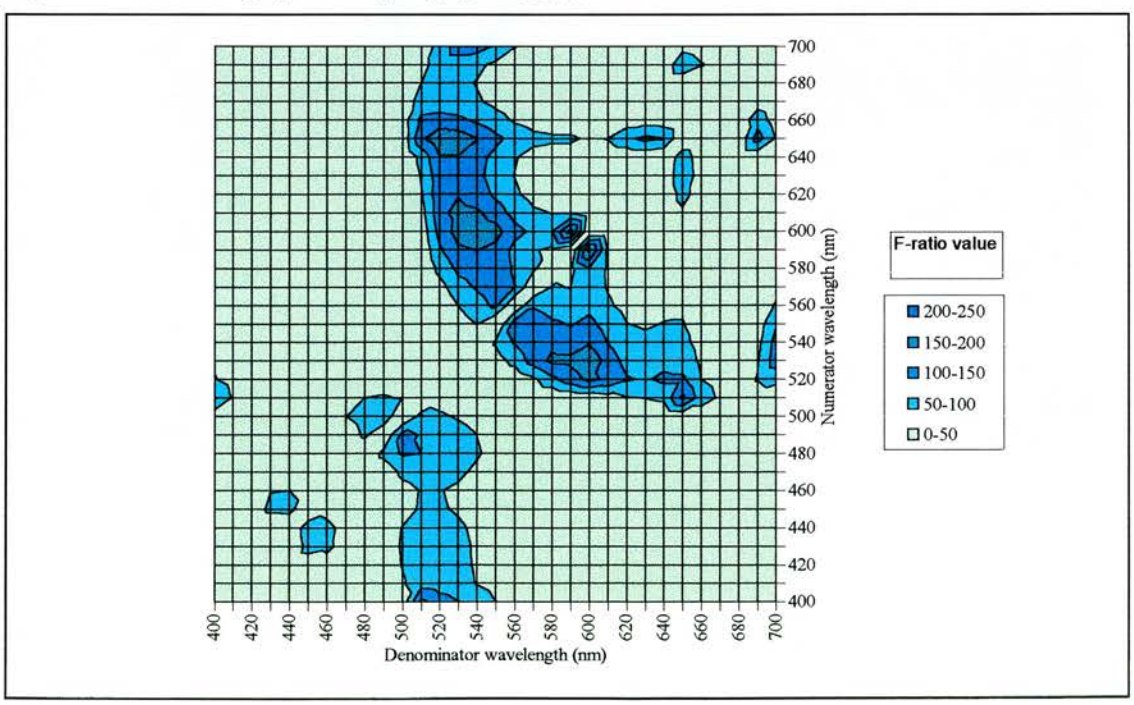

Figure 62. The Chlorophyte and Bacillariophyte $R_{PHY}(\lambda)$ F-ratio matrix.

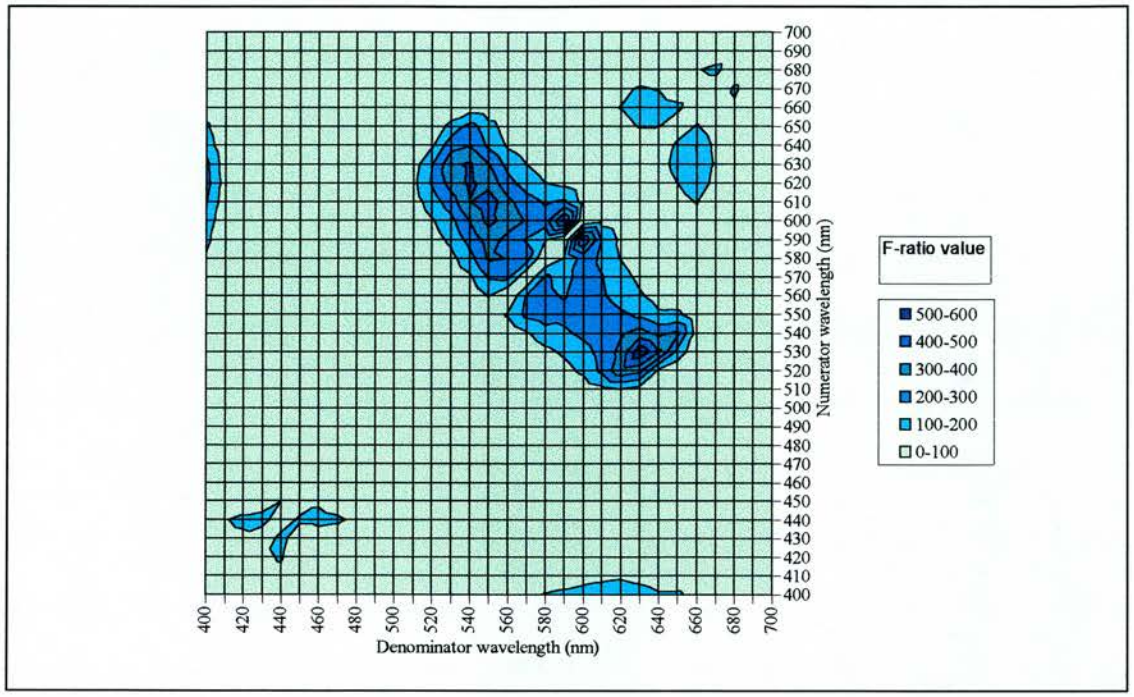

Figure 63. The Cyanophyte and Bacillariophyte $R_{PHY}(\lambda)$ F-ratio matrix.

The $R_{PHY}(\lambda)$ two-class F-ratio matrices have therefore confirmed the discriminatory power of the $a_{PHY}(\lambda)$ ratios but have also identified other $R_{PHY}(\lambda)$ ratios as being better in their ability to separate the various combinations of two classes. The wavelengths identified in the $a_{PHY}(\lambda)$ ratios were shown to be centres of absorption for class marker pigments. However there are no marker pigments reported in the literature as absorbing at the wavelengths identified by the strongest $R_{PHY}(\lambda)$ ratios. Consequently there is no simple explanation for the discriminatory strength of these $R_{PHY}(\lambda)$ ratios. There are however tentative explanations for this occurrence:

1. The strongest Chlorophyte and Cyanophyte $R_{PHY}(\lambda)$ ratio has both wavelengths located near the centre of the CHL*a* far-red absorption peak. Analysis of the $R_{PHY}(\lambda)$ spectra for these two classes shows that there is significant difference between the location of the CHL*a* $R_{PHY}(\lambda)$ minima at *ca.* 675 nm (Table 10). These minima are located towards the shorter wavelengths for the Chlorophytes (mean value of 669 nm) while the Cyanophytes have these reflectance minima located towards the longer wavelengths (with a mean at 679 nm). The locations of the Chlorophyte $R_{PHY}(\lambda)$ minima were possibly being influenced by the proximity of a CHL*b* absorption band to the main CHL*a* absorption band, the combination of the two centres of absorption resulting in a shift in the $R_{PHY}(\lambda)$ minima. Alternatively the Chlorophytes may have a form of CHL*a* which absorbs at shorter wavelengths than the Cyanophytes. ANOVA shows that the locations of these minima can be used to explain phytoplankton 'class' giving a significant F-ratio value of 238.3. It should be noted that there was no significant between-class difference in the location of the far-red $a_{PHY}(\lambda)$ peak for the phytoplankton species used in the experimental tank;

2. The strongest Chlorophyte and Bacillariophyte R_{PHY} ratio is 590 : 598 nm and the strongest Cyanophyte and Bacillariophyte reflectance ratio is at 588 : 600 nm. Bacillariophytes are reported as having CHL*c*, which is absent from both the Chlorophytes and Cyanophytes, and has a reported absorption band at 590 nm. This CHLc absorption feature is not obvious in the $a_{PHY}(\lambda)$. Furthermore there are no pigment absorption centres reported at *ca.* 600 nm for pigments occurring in either the Chlorophytes or Cyanophytes. Thus there is no marker pigment absorbing at the denominator wavelength for these ratios;
3. Ratios using these juxtaposed wavelengths are essentially recording the slope of the $R_{PHY}(\lambda)$. Each class displays a smooth, featureless, down-trending slope between the $R_{PHY}(\lambda)$ peak at *ca.* 550 nm and trough at *ca.* 675 nm. The gradient of this slope differs between the classes: in ascending order of steepness are the Bacillariophytes; the Chlorophytes; and the Cyanophytes. This spectra shape arises because of the presence of the strong 635 nm absorption band caused by PCC in the Cyanophytes. The PCC absorption causes a reduction in $R_{PHY}(\lambda)$ hence the Cyanophytes have the steepest downtrending slope. PCC is not present at high concentrations in the Chlorophytes or the Bacillariophytes and thus their slopes are gentler. The Chlorophytes have steeper slopes than the Bacillariophytes because their 540 nm R_{PHY} peak is considerably higher than the Bacillariophyte 560 nm R_{PHY} peak while both downtrend to the R_{PHY} minima at 675 nm;
4. The $R_{PHY}(\lambda)$ data set is restricted to six species from three classes at a wide range of concentrations while the $a_{PHY}(\lambda)$ data set included a broader range of species from each class. It is therefore possible that the results from the ANOVA test on the $R_{PHY}(\lambda)$ data set have been biased by the limitations of that data set;
5. It is possible that the assumption of scattering being non-trending with wavelength throughout the visible spectrum is incorrect. If anomalous diffraction (Van de Hulst 1956) is influencing the scattering spectra in the region of an absorption band this might influence the location of the $R_{PHY}(\lambda)$ minima especially at the high cellular concentrations used in the experimental tank. A scattering influence on $R_{PHY}(\lambda)$ would not detract from an explanation based upon marker pigment absorption.

Further analysis would be required to explain the lack of conformity between the $a_{PHY}(\lambda)$ and $R_{PHY}(\lambda)$ ratios. However the analytical interpretation for the $a_{PHY}(\lambda)$ ratios and the significant discriminatory strength of these ratios when applied to reflectance spectra suggests that they are the most applicable ratios for two-class discriminations of $R_{PHY}(\lambda)$.

7.1.4.2 Using ratios to classify the phytoplanktonic composition of natural water bodies

The class identification ratios require certain modifications if they are to be used in a predictive capacity. As was discussed in the previous chapter CHL*a* prediction ratios have frequently been used and these ratios tend to be lake-specific to accommodate the optical properties of each natural water body (Dekker and Peters 1993, George and Charlton 1996). Consequently, even though the class

discriminating ratios have been developed in a semi-analytical manner, there is also likely to be a requirement for adjustments in the locations of the ratio wavelengths to accommodate the individual optical characteristics of each water bodies.

To exploit the predictive capability of these class identification ratios they require modification into working algorithms. There are a number of ways in which these two-class ratios could be converted into class identification routines. Absolute predictions of cell numbers by class would require knowledge of how changes in class cell concentrations are reflected in both the absolute value and the overall distribution of the results from each ratio. Given the limitations of the set of phytoplankton reflectance spectra and the restrictions placed on those $R_{PHY}(\lambda)$ by the experimental tank, it is likely that such distributions will not be applicable to natural water bodies. Were a fuller set of $R_{PHY}(\lambda)$ available then the absolute response of a particular ratio to changes in class concentration could be used to classify, on the basis of proximity to the nearest mean, ratio values for each class. Alternatively, full use of the distribution information could be made to assign spectra on the basis of maximum likelihood of class membership. The distribution of values for the two class ratios for both of the classes in the ratio can be converted into likelihoods using the following equation:

$$L = \frac{1}{\sqrt{2\pi\sigma}} \exp\left[-(x-\mu)^2 \Big/ (2\sigma^2) \right] \quad \text{.....7.1.4.2.1}$$

where L is the likelihood of class membership (ranging 0 to 1), σ and μ are the standard deviation and mean of the ratio values for the particular class under consideration, x is the ratio value for the sample spectra (Hope 1968). L would be calculated with respect to both of the phytoplankton classes and assignment made according to which ever class had the highest L value.

The advantage of a maximum likelihood classification based upon *a priori* knowledge of distributions is that they could be used to obtain a 'fuzzy' rather than a hard classification (Foody 1996). Natural phytoplankton populations tend to be composed of more than one class and although a 'hard' classification of the dominant class would undoubtedly be useful to water managers, it would be more useful if some idea of the relative concentration of each class could be ascertained. A fuzzy classification reporting the likelihood of the presence of each class would provide some indication of the relative concentrations of the classes present in a sample. A drawback of classifications based upon *a priori* knowledge of distributions is that they are limited by the data set from which these distributions were calculated (Foody 1996). Furthermore outlying data points in the distribution data set can severely reduce the accuracy of a classification.

No application of the class ratios to the test data sets is presented here. What could be achieved with the class ratios at this stage is a relative assessment of time-series reflectance spectra in terms of determining the relative changes in the dominant class of a phytoplankton crop over time. This technique does not use the absolute value of any ratio, but merely utilises the relative changes in the

ratio as crop composition changes. This relative assessment has been employed to classify the Esthwaite Water reflectance spectra presented in Chapter Five.

Two class ratios will only distinguish between the two classes for which the ratio was developed. Attempts were made to develop multi-class ratios (three or more classes). The effective two class ratios used observations centred upon the absorption centres of pigments which were present in one of the classes and absent from the other, hence maximising the discrimination in the resulting ratio. Restricting the analysis to a simple ratio and including a third class meant that there was one less observation than there were classes to discriminate. Consequently the *Fratio.pas* programme identified ratios centred on marker pigment absorption for two of the classes, the ratio value for the third class merely occupied values lying between the extreme values other two classes. As such a default discrimination arose because the third class had neither of the two marker pigments, so the ratio values for that class would exhibit little variation. Any class which possessed neither of the two class marker pigments would experience this default classification. Thus no further analysis was made of multi-class two-band ratios.

Restricting the analysis to two class ratios has the drawback of requiring a choice to be made of which particular two class ratio is most applicable to a particular water body at a particular time of year. For instance there may be a crop of Cyanophytes in a lake, the dominance of which will be visible in two-class ratios which include Cyanophytes. However those two-class ratios which do not include Cyanophytes will still provide an output value which will indicate that one or other of the classes in the ratio is dominant when this is manifestly not the case. It would therefore be inadvisable to apply all the two class ratios to a spectra from a target water body as this may lead to conflicting results. What is required is an informed choice of which ratios are most applicable. Thus some form of expert system may be required to assist operators in their choice of ratio. Such a systems could be based on *a priori* knowledge of the likely state of a water body at a particular time in the year utilising the work on phytoplankton succession of Reynolds (1980). This idea is addressed in the concluding chapter as it is relevant to all the proposed identification routines.

7.1.5 Summary of class identification ratios

The two class ratios introduced in this section have highlighted the importance of a small number of class marker pigments to the problem of phytoplankton identification. Through the absorption properties of these marker pigments it has been possible to develop ratios which vary in accordance with class concentrations. It should not prove difficult to map class distributions were a spatial data set with observations at the particular ratio wavelengths available, although more work on developing the identification algorithms for individual water bodies is required.

7.2 Canonical discriminant analysis based classifications

Discriminant analysis has been widely used as a land cover classification technique in remote sensing (Foody 1996). Jakubauskas (1996) used canonical correlation analysis to analyse the relationship between forest canopy type and reflectance. In the aquatic sciences Johnsen *et al.* (1994) used discriminant analysis to classify phytoplankton classes from *in vivo* absorption spectra. This section draws on the work of Johnsen *et al.* (1994) using canonical discriminant analysis to classify phytoplankton classes from their absorption spectra, extending the technique to reflectance spectra.

7.2.1 Introduction to canonical discriminant analysis

Discriminant analysis and canonical correlation analysis are both instances of regression on a set of two or more variables (Hope 1968). Both techniques seek to maximise the correlation coefficient associated with a regression. Discriminant analysis assigns individuals to groups with the minimum level of misclassification, consequently it allows analysis of the decision rules for allocating new individuals to groups and sheds light upon the optimum set of variables required for a successful classification (Hope 1968, Kendall 1968). It parsimoniously, but effectively, describes the measured differences between distinctly separate groups by projecting them into a suitable sub-space by transforming multi-variate data into a single discriminant score (Cooley and Lohnes 1971). Discriminant analysis therefore identifies real differences in the multivariate input data for each group, separating them significantly enough to minimise the chance of misclassification (Kendall 1968). This between-group separation is achieved through the calculation of orthogonal discriminant functions. The maximum number of discriminant functions is one less than the smaller of either the number of groups or the number of variables (Hope 1968).

It is possible to compute a discriminant analysis as a special case of canonical correlation analysis in which a set of dummy variables carries the group membership information (Cooley and Lohnes 1971). Canonical correlation analysis is a dimension reduction technique originally developed by Hotelling (1935, 1936) for analysing the relationship between two sets of variables. Given two sets of variables canonical correlation analysis finds a linear combination from each set, called a canonical variable, such that the correlation between the two canonical variables is maximised. This correlation is the first canonical correlation. The weighting coefficients for the linear combination of the sets are called canonical coefficients. Having ascertained the first canonical correlation it may be possible to identify further canonical correlations, with different canonical coefficients, subject to the restriction that the new canonical correlations are uncorrelated with all other canonical correlations (Cooley and Lohnes 1971). The canonical functions are not orthogonal (Hope 1968).

Because canonical discriminant analysis is a dimension reduction technique it provides a means by which the intercorrelations within two data sets can be accounted for while simultaneously examining the relations between those data sets (Jakubauskas 1996). When applied to the problem of identifying

the dominant phytoplankton class in hyperspectral reflectance spectra it allows the spectra to be reduced to a the minimum number of observations (*i.e.* the minimum number of wavebands) that are necessary to discriminate between the phytoplankton classes. Reducing the number of observations is important because of the limitations on the number of bands available when using traditional airborne remote sensors with spatial capabilities.

7.2.2 Canonical discriminant analysis methods

To ascertain the minimum number and location of wavebands necessary to distinguish between phytoplankton class a stepwise discriminant analysis was performed using SPSS v6.0. The stepwise selection rule was 'Minimise Wilks' Lambda'. Wilks' Lambda is a multivariate statistic for evaluating whether or not two, or more, groups come from populations with the same mean. It ranges from zero to one, with smaller values suggesting a significant difference between-group means. Hence the wavelengths with the lowest Wilks' Lambda should best differentiate between phytoplankton classes. The selection rule also ensured that wavelengths with F-ratio values < 1.0 were not selected and if an already selected wavelength's F-ratio value fell below 0.5 when a new wavelength was included then it was dropped from the selection. Furthermore, the minimum tolerance was set to 0.001 (tolerance describes the linear relationships between wavelengths and is calculated according to $1 - R^2$ when absorption at one wavelength is used to predict absorption at another wavelength). The number of discriminant functions required was restricted to two.

The 'Minimise Wilks Lambda' selection rule was chosen as the stepwise selection rule following analysis of the various options available on SPSS. The other selection rules were: 'Minimise the maximum F-value' whereby the variable that maximises the smallest F-ratio between pairs of groups is entered; 'Maximise the Rao's V statistic' whereby the variable which produces the largest increase in Rao's V statistic is entered; and 'Minimise the sum of the residuals' whereby the variable that minimises the sum of the unexplained variation for all pairs of groups is entered. The stepwise selected wavelengths were assessed by analysis of the multivariate statistics used to describe differences between-group means when each selection rule was on a sub-set of the pure culture $a_{PHY}(\lambda)$ database. The results of this precursory analysis highlighted the 'Minimise Wilks' Lambda' selection rule as being the most successful in selecting those wavelengths suitable for between-class discrimination.

The input data set for the stepwise analysis consisted of 88 pure culture $a_{PHY}(\lambda)$ reduced to 2 nm intervals from 400 to 710 nm (34 Chlorophytes, 23 Cyanophytes, 17 Bacillariophytes, 8 Cryptophytes and 6 Euglenophytes). Because $a_{PHY}(\lambda)$ change spectrally not only with phytoplankton class composition but also with biomass it was necessary to transform the $a_{PHY}(\lambda)$ to both normalise the spectra and to stabilise the within spectra variance caused by concentration effects across the wavelengths. These transformations were: natural logged absorption coefficients, $\ln(a_{PHY}(\lambda))$;

chlorophyll-*a* specific absorption coefficients, $a_{PHY}^*(\lambda)$; and natural logged specific absorption coefficients, $\ln(a_{PHY}^*(\lambda))$. In view of the low number of Cryptophyte and Euglenophyte $a_{PHY}(\lambda)$ the stepwise procedure was run for each transformation firstly on the full data set and secondly on a data set excluding these two classes.

Having selected the appropriate wavelengths using the SPSS stepwise discriminant analysis a canonical discriminant analysis (CANDISC) was performed on SAS v6.07 (SAS Institute Inc. Cary, USA) running on a Unix based Sun Sparc 20. The CANDISC routine was used to assess the discriminatory capacity of the various transformations and wavelength selections.

7.2.3 Canonical discriminant analysis results

The stepwise procedure selected the $a_{PHY}(\lambda)$ wavelengths for each transformation as shown in Table 24. These results imply that only a few wavelengths were necessary for the identification of phytoplankton class. However the selection of wavelengths depended upon the data transformation and the number of classes under investigation. For the $\ln(a_{PHY}^*(\lambda))$ data set thirteen $a_{PHY}(\lambda)$ wavelengths were included in the stepwise analysis of the five classes, and ten wavelengths for the three classes, five wavelengths were common to both. For the $\ln(a_{PHY}(\lambda))$ data set twelve $a_{PHY}(\lambda)$ wavelengths were included for the five classes and nine for the three classes, three wavelengths were common to both. For the $a_{PHY}^*(\lambda)$ transformation only seven wavelengths were required for the five classes and eight for the three classes, only two wavelengths were common to both.

There were similarities in the wavelengths selected for the different data transformations. For the five class data set wavelengths 636 ± 2 nm and 708 ± 2 nm were included by the stepwise analysis of the $\ln(a_{PHY}(\lambda))$ and $\ln(a_{PHY}^*(\lambda))$ data set and 654 ± 2 nm and 700 ± 4 nm were included in all three transformations. For the three class data set 536 ± 2 nm, 668 ± 2 nm and 682 ± 2 nm were included in the stepwise analysis for the $\ln(a_{PHY}(\lambda))$ and $\ln(a_{PHY}^*(\lambda))$ data sets; 608 ± 2 nm and 614 nm were included in the $\ln(a_{PHY}^*(\lambda))$ and $a_{PHY}^*(\lambda)$ data sets. 652 ± 2 nm and 660 ± 2 nm were common to all three transformations.

Table 24. The wavelengths (nm) selected by the 'Minimise Wilks' Lambda' stepwise selection rule for the five class (Chlorophyte, Cyanophyte, Bacillariophyte, Cryptophyte and Euglenophyte) and three class (Chlorophyte, Cyanophyte and Bacillariophyte) data sets under various transformations. A negative sign indicates that a wavelength was de-selected during that step.

Data set	*Step 1*	*Step 2*	*Step 3*	*Step 4*	*Step 5*	*Step 6*	*Step 7*	*Step 8*	*Step 9*	*Step 10*	*Step 11*	*Step 12*	*Step 13*
Five classes													
$\ln(a_{PHY}^*(\lambda))$	560	606	654	574	636	698	514	594	614	530	544	702	708
$a_{PHY}^*(\lambda)$	564	600	704	484	654	674	686						
$\ln(a_{PHY}(\lambda))$	652	636	550	570	612	656	698	512	704	710	490	432	
Three classes													
$\ln(a_{PHY}^*(\lambda))$	608	654	574	538	668	614	682	660	464	594			
$a_{PHY}^*(\lambda)$	610	512	484	544	480	658	652	588	-544	614			
$\ln(a_{PHY}(\lambda))$	652	636	544	670	626	536	684	660	428				

The discriminatory strength of the wavelengths selected for each transformation was assessed using the CANDISC procedure. Discriminatory strength was measured in terms of the Wilks' Lambda statistic and the size of each canonical correlation. The smaller the Wilks' Lambda (from 0 to 1) and the larger the canonical correlation (from 0 to 1) the stronger the relationship between the variables generating that correlation, consequently the better the canonical discrimination. The various Wilks' Lambda values and canonical correlations are shown in Table 25.

Table 25. The Wilk's Lambda and canonical correlations for five and three class data sets with the various transformations described in the text, using the wavelength given in Table 24. All Wilks' Lambda statistics are significant at > 99.9% level of probability.

Data set	*Wilks' Lambda*	*First canonical correlation*	*Second canonical correlation*
Five classes			
$\ln(a_{PHY}^*(\lambda))$	0.001	0.981	0.918
$a_{PHY}^*(\lambda)$	0.058	0.861	0.781
$\ln(a_{PHY}(\lambda))$	0.002	0.979	0.913
Three classes			
$\ln(a_{PHY}^*(\lambda))$	0.004	0.984	0.922
$a_{PHY}^*(\lambda)$	0.070	0.894	0.808
$\ln(a_{PHY}(\lambda))$	0.009	0.972	0.916

The strength of the Wilks' Lambda (Table 25) suggests that for both the five class and three class data sets the wavelengths selected by the stepwise analysis for the $\ln(a_{PHY}^*(\lambda))$ transformations (shown in Table 24) gave the strongest between-class discrimination. Similarly those wavelengths selected by the $a_{PHY}^*(\lambda)$ transformations produced the weakest discrimination.

The canonical discrimination statistics shown in Table 25 for each group of stepwise selected wavelengths for the five and three class data sets subjected to the same transformations suggests that there is no reduction in discrimination when using the larger data set. However graphical representation of the canonical scores show, that for the same transformation the three classes are better discriminated than the five classes. Figure 64 and Figure 65 show the canonical scores plotted on the two canonical discrimination functions for the stepwise selected wavelengths for the five class

$\ln(a_{PHY}^*(\lambda))$ data set, and the three class $\ln(a_{PHY}^*(\lambda))$ data set respectively. The wavelengths selected by the five class stepwise discrimination analysis failed to adequately separate the Euglenophytes and the Chlorophytes. This result concurs with the findings of Johnsen *et al.* (1994) who also found difficulty in separating these phytoplankton classes.

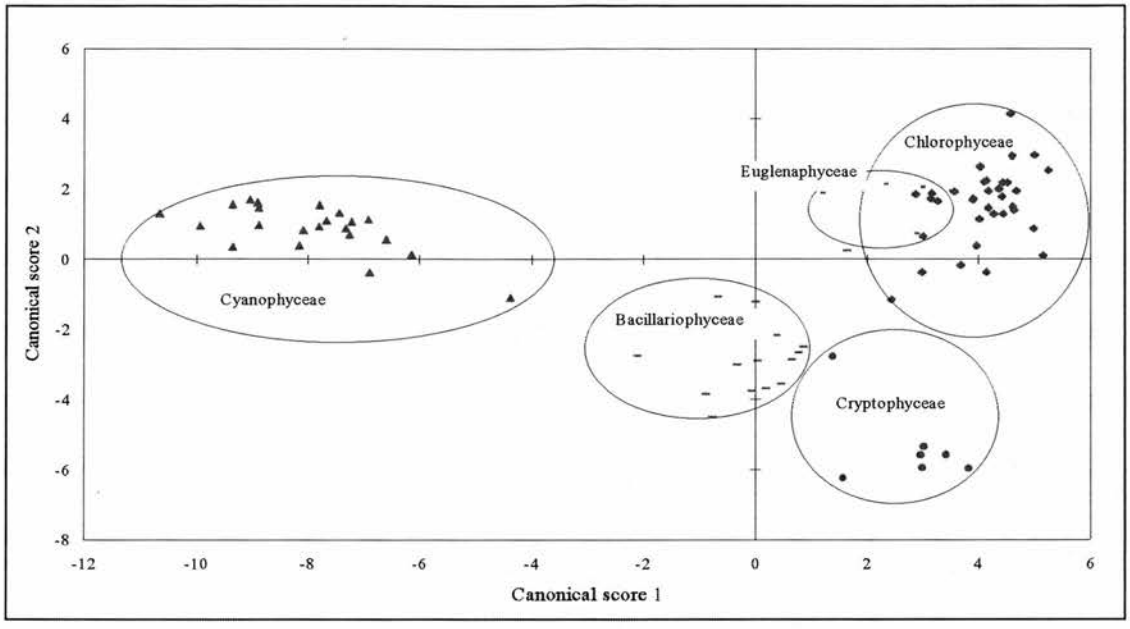

Figure 64. The standardised canonical scores for the stepwise selected wavelengths from the five class $\ln(a_{PHY}^*(\lambda))$ data set. The ellipses have been added by hand to highlight particular class groupings.

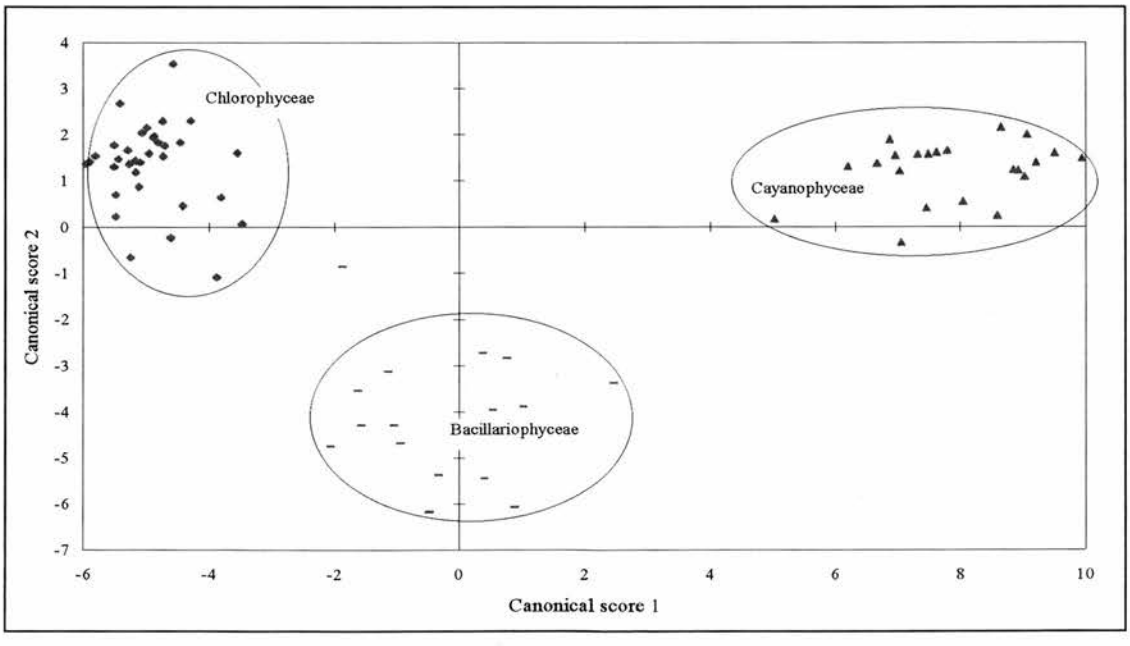

Figure 65. The standardised canonical scores for the stepwise selected wavelengths from the three class $\ln(a_{PHY}^*(\lambda))$ data set. The ellipses have been added by hand to highlight particular class groupings.

The stepwise routine was performed with the selection criteria set to include more wavelengths than the absolute minimum. Johnsen *et al.* (1994) found that only five wavelengths were necessary for the discrimination of ten classes. This compares to the thirteen wavelengths selected for the canonical discrimination of five classes and ten wavelengths for a three class discrimination presented here. Study of the raw canonical coefficients suggest, that there will be only a minor reduction in the discrimination capacity if some of the more redundant wavelengths are removed from the analysis. The magnitude of the raw canonical coefficients relative to the magnitude of the input data, and the sign of the raw canonical coefficients on each canonical variate, indicates the role played by each wavelength in the discrimination. Of the stepwise selected wavelengths shown in Table 24 for the five class canonical discrimination, it was inferred that 544, 636, 698, 702 and 708 nm had relatively high redundancy, while 530, 574, 594, 606, 614 and 654 nm were all strongly involved in the discrimination. For the three class $\ln(a_{PHY}^*(\lambda))$ transformation the only redundant wavelength was 464 nm. Optimising the stepwise wavelengths did not perceptibly reduce the within-class clustering or the between-class separation.

7.2.4 Canonical discriminant analysis discussion

It has been shown that the most successful stepwise discrimination analysis was for the $\ln(a_{PHY}^*(\lambda))$ transformed pure culture $a_{PHY}(\lambda)$. This result concurs with the transformation used by Johnsen *et al.* (1994). Calculating $a_{PHY}^*(\lambda)$ accounts for the variation in species $a_{PHY}(\lambda)$ attributable to cell biomass (§ 2.1.5.2) and calculating $\ln(a_{PHY}(\lambda))$ gives weight to the weaker spectral absorption regions of the $a_{PHY}(\lambda)$. Combining the two achieves the most appropriate normalisation.

Although it is not possible to identify which individual wavelengths are responsible for the discrimination of each class, it is possible to identify those wavelengths which are located at wavelengths where pigment absorption peaks as reported in the literature (Appendix Two). The pigments absorbing at the wavelengths selected for the five and three class data sets under the $\ln(a_{PHY}^*(\lambda))$ transformation are shown in Table 26, along with the class affiliations of the pigments.

Table 26. Pigments with absorption features reported to be located at ± 5 nm of the optimised stepwise selected wavelengths for the five and three class $\ln(a_{PHY}^*(\lambda))$ data set. Pigments in bold are those most likely to be responsible for the inclusion of that wavelength by the stepwise procedure. The class associations of these pigments are indicated in superscript.

Five class data set wavelengths (nm)	*Pigments*	*Three class data set wavelengths (nm)*	*Pigments*
530	**FCX**[3,4]	538	**FCX**[3], **PCE**[2], SPX[1]
544	**PCE**[2,4], **FCX**[3,4]	574	**PCE**[2]
560	**PCE**[2,4], PCC[2,4]	608	PCC[2]
574	**PCE**[2,4]	614	**PCC**[2]
594	PCE[2,4], CART[2,3,4,5]	654	**APC**[2], **CHL*b***[1], PCE[2]
606	PCE[2,4], PCC[2,4]	660	**CHL*b***[1], CHL*a*[1,2,3]
614	**PCC**[2,4]	668	CHL*a*[1,2,3], **APC**[2], **CHL*b***[1]
636	**CHL*c***[3,4]	682	CHL*a*[1,2,3]
654	**APC**[2,4], **CHL*b***[1,5], PCE[2,4]		

[1] = Chlorophytes, [2] = Cyanophytes, [3] = Bacillariophytes, [4] = Cryptophytes and [5] = Euglenophytes

Table 26 shows that the absorption by PCE and PCC has been identified at various wavelengths. These biliproteins are marker pigments for both the Cyanophytes and Cryptophytes. Several types of PCE and PCC can be distinguished spectroscopically (§ 4.1.2). Furthermore both are likely to exhibit distinct differences in spectral absorption when samples from different classes are compared. The relative quantities of these biliproteins are also likely to change through photoadaptation (Britton 1983, Ganff *et al.* 1991). The inclusion of wavelengths where absorption by CHL*a* dominates suggests that either the normalisation for biomass has not been completely effective, or that the canonical discriminant analysis is exploiting nuances in the between-class absorption characteristics of CHL*a*.

The stepwise selection procedure has therefore empirically reduced the hyperspectral $a_{PHY}(\lambda)$ to a set of taxonomically important wavelengths located at the absorption centres of class marker pigments. However the $a_{PHY}(\lambda)$ data set upon which the stepwise analysis performed this reduction was restricted to 18 species from five classes, it is therefore possible that given a different set of input species $a_{PHY}(\lambda)$ the stepwise routine would have selected different wavelengths. To compile a more universally applicable set of wavelengths the pigment database was consulted to ascertain the wavelengths of the appropriate class marker pigments. These analytically selected wavelengths and the results of applying them from the five class $\ln(a_{PHY}^*(\lambda))$ data set to CANDISC are shown in Table 27. The canonical scores show reasonable within-class clustering and between-class separation (Figure 66). The analytically selected wavelengths achieved a Wilks' Lambda of 0.004, with the first two canonical correlations being 0.973 and 0.969. These compare well to the canonical discrimination statistics for the stepwise selected wavelengths for the same data set (Wilks' Lambda of 0.001 and canonical correlations of 0.981 and 0.918). There is some overlap between the analytically selected wavelengths and the stepwise selected wavelengths with 530, 560, 630 (a surrogate for 636) and 654 nm being common to both selections.

Table 27. The raw canonical coefficients and pigment associations for the analytically selected wavelengths for the five class $\ln(a_{PHY}^*(\lambda))$ data set. The class associations of these pigments are indicated. The linear summation of the product of the input variable and the appropriate canonical coefficient plus the constant give the standardised canonical scores shown in Figure 66.

Wavelength (nm)	*Canonical coefficient 1*	*Canonical coefficient 2*	*Associated pigments*
454	9.915	-9.404	CHL$c^{3,4}$, CHL$b^{1,5}$
464	-7.529	28.764	CHL$c^{3,4}$, DDX
474	-23.761	-14.081	CHL$b^{1,5}$
490	28.587	-9.684	DDX^{5}
502	-7.382	2.423	MYX^{2}
530	1.071	1.728	FCX^{3,4}
560	-2.891	2.287	PCE^{2,4}, PCC^{2,4}
630	15.518	-2.821	PCC^{2,4}
656	-13.590	1.058	CHL$b^{1,5}$
Constant	0.904	3.444	

[1] = Chlorophytes, [2] = Cyanophytes, [3] = Bacillariophytes, [4] = Cryptophytes and [5] = Euglenophytes

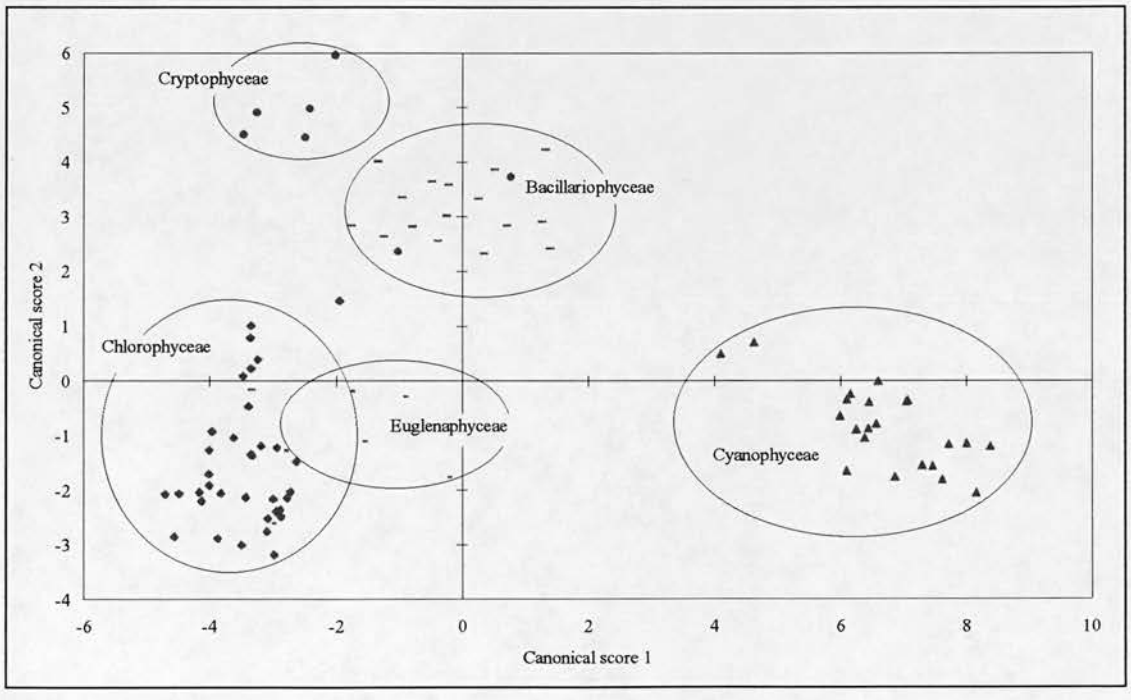

Figure 66. The standardised canonical scores for the analytically derived $\ln(a_{PHY}^*(\lambda))$ wavelengths for the five classes used in this study (Table 27). The ellipses have been added by hand to highlight particular class groupings.

Application of canonical discriminant analysis to the problem of identifying the dominant class of $a_{PHY}(\lambda)$, and therefore of $R_{PHY}(\lambda)$, will be most successful for spectra from natural Case I water bodies, if the analytically selected wavelengths are used. This is because the presence of other optically active components will interfere with the discriminant analysis as they were not present in the spectra used to develop the canonical coefficients.

Having ascertained that canonical discriminant analysis can be used to identify class from $a_{PHY}(\lambda)$ it is necessary to transpose the technique to $R_{PHY}(\lambda)$. For this purpose the stepwise procedure used to ascertain the minimum number of wavelengths necessary for effective discrimination of the $R_{PHY}(\lambda)$ spectra under the following transformations: the raw reflectance spectra, $R_{PHY}(\lambda)$; the natural-log of reflectance, $\ln(R_{PHY}(\lambda))$; and the natural-log of the 'specific' reflectance spectra which has been 'corrected' for CHL*a* concentration, $\ln(R_{PHY}^*(\lambda))$. The results of this stepwise discriminant analysis are given in Table 28.

Table 28. The wavelengths (nm) selected by the 'Minimise Wilks Lambda' stepwise rule for the discrimination of three classes for the three data transformations applied to the pure culture phytoplankton reflectance spectra.

Data set	*Step 1*	*Step 2*	*Step 3*	*Step 4*	*Step 5*	*Step 6*	*Step 7*
$R_{PHY}^*(\lambda)$	400	596	526	570	720	646	680
$\ln(R_{PHY}^*(\lambda))$	652	632	544	520			
$R_{PHY}(\lambda)$	400	640	522	720	664	570	

Seven reflectance wavelengths were selected for the $R_{PHY}^*(\lambda)$ data set, six for the $R_{PHY}(\lambda)$ data set and four for the $\ln(R_{PHY}^*(\lambda))$ data set. Certain wavelengths were selected by the step-wise procedure for more than one transformation. The reflectance wavelengths 400, 570, 644 ± 4 and 720 nm were selected by the stepwise analysis for the both $R_{PHY}^*(\lambda)$ and $R_{PHY}(\lambda)$ transformation. 522 ± 4 nm was selected for all the transformations.

Canonical discriminant analysis showed that the wavelengths selected for the $R_{PHY}^*(\lambda)$ transformation produced the smallest Wilks' Lambda (0.004) and the strongest canonical correlations (0.975 and 0.956) of the three transformations. Furthermore plots of the canonical scores for the $R_{PHY}^*(\lambda)$ transformation selected wavelengths showed the strongest propensity for within-class grouping and between-class separation, using both canonical variates to effectively separate the phytoplankton classes. It can therefore be concluded that the most effective canonical discriminant analysis is being performed upon those wavelengths selected by the stepwise analysis from the $R_{PHY}^*(\lambda)$ transformed data set. The between-class separation afforded by the $R_{PHY}^*(\lambda)$ transformation at the selected wavelengths is shown in Figure 67.

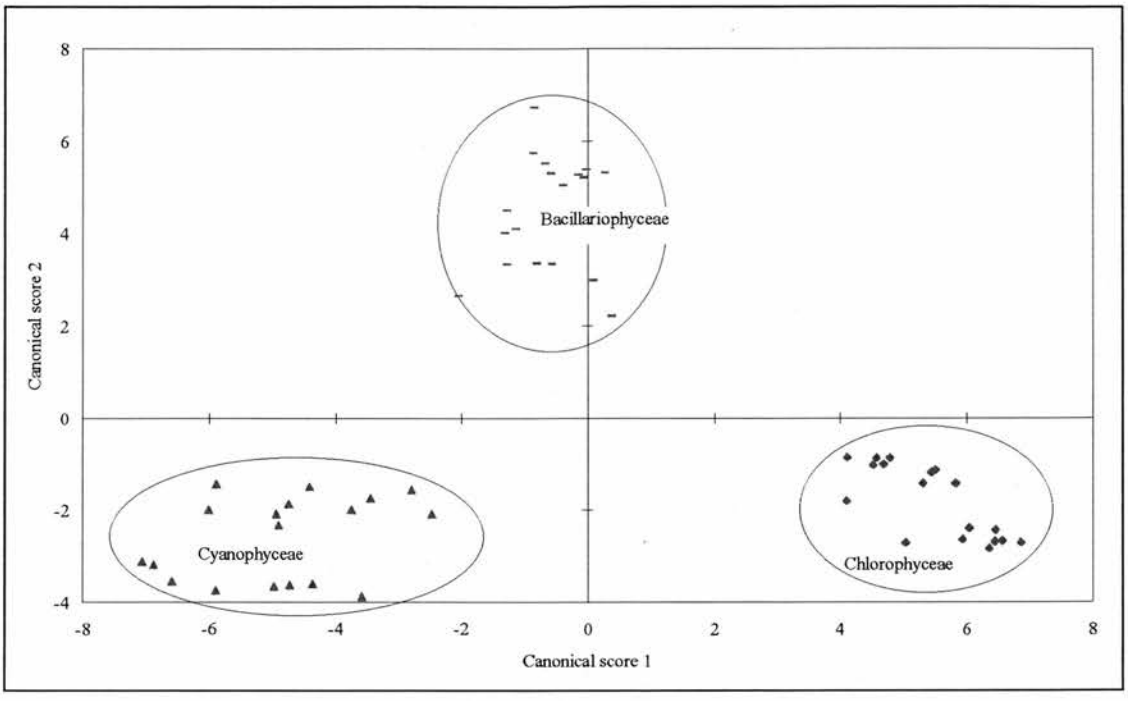

Figure 67. The standardised canonical scores for the canonical discriminant analysis performed upon the stepwise selected wavelengths for the $R_{PHY}^*(\lambda)$ data set. The ellipses have been added by hand to highlight particular class groupings.

Again the stepwise discriminant analysis has selected wavelengths where taxonomically important pigments are absorbing. However there are two wavelengths from either end of the input reflectance spectra which have no pigment association (400 and 720 nm). The inclusion of these two wavelengths at either end of the spectra most likely serves to remove the influence of any trending scattering spectra.

The stepwise discriminant analysis described above acts as a comparison for assessing the performance of the analytical wavelength set (Table 27) when applied to $R_{PHY}(\lambda)$. Using the analytical wavelengths for the pure culture $R_{PHY}^*(\lambda)$ for input into CANDISC gave a Wilks' Lambda of 0.003 and canonical correlations of 0.979 and 0.965, both of which were significant and stronger than the canonical correlations for the $R_{PHY}^*(\lambda)$ selected wavelengths. The canonical scores for these reflectance wavelengths can be seen in Figure 68. The between-class separation afforded by these wavelengths for the $R_{PHY}^*(\lambda)$ data set is as good as the stepwise selected wavelengths for the same data set (Figure 67).

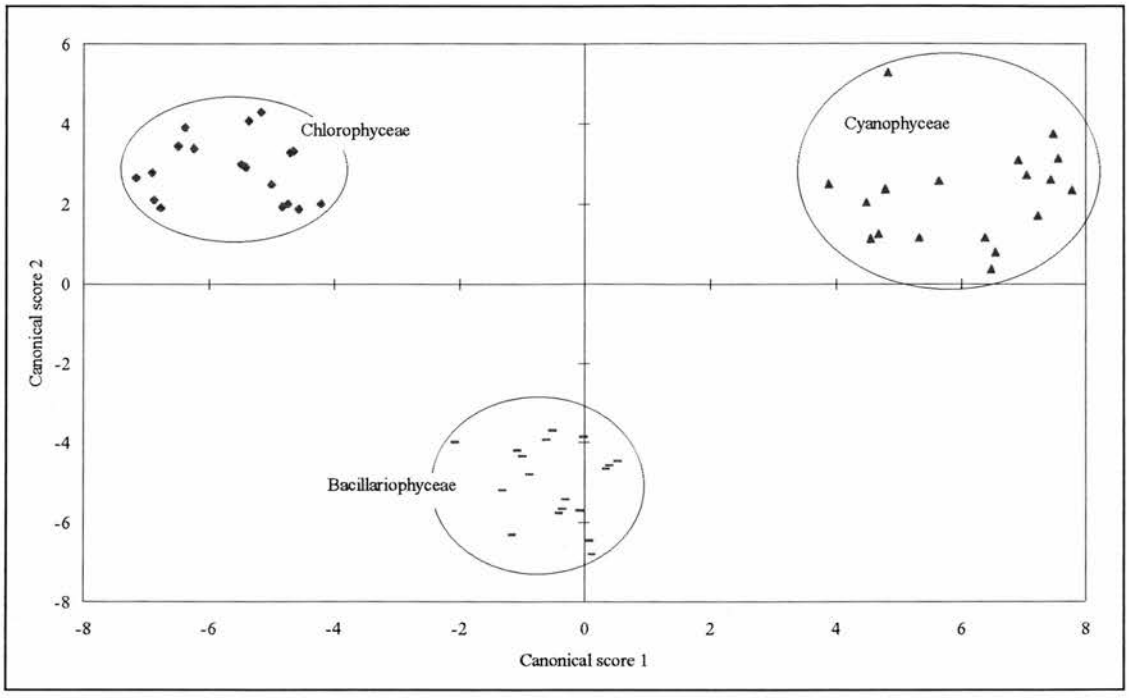

Figure 68. The standardised canonical scores for the $R_{PHY}^*(\lambda)$ data set using the analytically derived wavelength set based on the canonical discriminant analysis of the absorption spectra (Table 27). The ellipses have been added by hand to highlight particular class groupings.

The stepwise selection of the $R_{PHY}^*(\lambda)$ data set included two end wavelengths. Inclusion of 400 and 720 nm to the analytically derived wavelengths greatly improved the canonical discriminant analysis for the $R_{PHY}^*(\lambda)$ data set achieving a Wilks' Lambda of 0.0008 and canonical correlations of 0.989 and 0.981. The canonical scores for these extended analytical $R_{PHY}^*(\lambda)$ wavelengths are shown in Figure 69 and the canonical coefficients in Table 29.

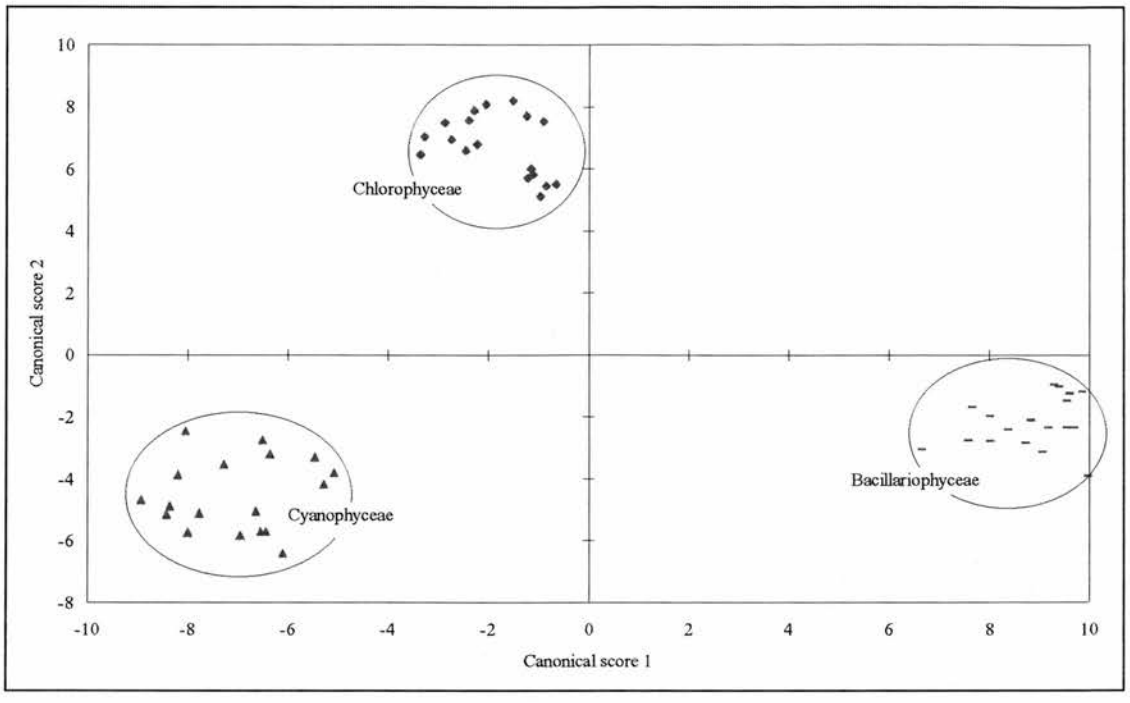

Figure 69. The standardised canonical scores for the extended analytically derived wavelengths (the wavelengths given in Table 27 plus 400 and 720 nm) applied to the $R_{PHY}^*(\lambda)$ data set. The ellipses have been added by hand to highlight particular class groupings. Note the improved within-class clustering and between-class separation compared with Figure 68.

Table 29. The canonical coefficients for the analytical wavelength set (extended to include two wavelengths at either end of the visible spectrum) using the $R_{PHY}^*(\lambda)$ transformation for the pure culture phytoplankton reflectance spectra. The resulting canonical scores are shown in Figure 69.

Wavelength (nm)	*Canonical coefficient for first canonical variable*	*Canonical coefficient for second canonical variable*
400	2125.8	-1134.5
454	-621.9	-140.0
464	-667.5	1035.5
474	1934.7	-432.4
490	-177.4	-434.3
502	-1173.5	-46.6
530	182.3	821.8
560	-798.8	-599.6
630	1930.7	192.4
656	-320.9	-160.2
720	-875.5	168.8
Constant	6.1	2.7

All the canonical discriminant analyses performed have used narrow band input data from hyperspectral data sets. This form of data is not always available to remote sensing, more typically remote sensors provide broader band data. It is important to ensure that the canonical discriminant technique performs adequately using broader bandwidths. Applying the analytical wavelength set at ± 4 nm with $\ln(a_{PHY}^*(\lambda))$ five class data set to CANDISC achieved a canonical discriminant analysis of

0.004, with the first two canonical correlations being 0.974 and 0.911. This compares well to the narrow band wavelengths for the same data set with no apparent loss in class discrimination. A similar result was found for the extended analytical wavelength set at ± 4 nm applied to the $R_{PHY}^*(\lambda)$ data set. Hence increasing the bandwidth does not adversely affect the canonical discriminant analysis.

Canonical correlations are linear combinations of canonical coefficients chosen to maximise the correlation with the canonical variables. For this study the number of canonical correlations has been restricted to two. The first canonical correlation will always explain a greater percentage of the variation in the input variables than will the second canonical correlation. Canonical discriminant analysis uses these canonical correlations to discriminate between input classes. Consider the two canonical variates for the analytical $\ln(a_{PHY}(\lambda)^*)$ wavelength set (Figure 66), the Cyanophytes class separation is being afforded by the first canonical variate which explains 75% of the variance in the input data, the other four classes have essentially the same scores on the first canonical variate with the between-class separation being afforded by the second canonical variate which only explains an extra 21% of the variance. It is not surprising therefore to find that for the first canonical correlation the strongest correlation between a single input variable and the resulting canonical variable is 0.373 for $a_{PHY}(630)$ where PCC absorbs which is the principal marker pigment for the Cyanophytes. (Note the absolute value of the coefficient indicates the relative contribution of a particular wavelength to the canonical variate (Jakubauskas 1996)). On the second canonical correlation the strongest correlation between the input wavelength variables and the resulting canonical variable is for $a_{PHY}(530)$ where FCX absorbs (correlation = 0.434) which is a marker pigment distinguishing between both the Chlorophytes and Euglenophytes and the Cryptophytes and Bacillariophytes.

Considering the extended analytical input wavelength set for the $R_{PHY}^*(\lambda)$ transformation (Figure 69 and Table 29) the strongest correlations between the $R_{PHY}^*(\lambda)$ input wavelengths and the first canonical variate are 400, 530, 560 and 630 nm (-0.162, -0.189, -0.154 and 0.146 respectively) indicating the importance of the end wavelength 400 nm, and the marker pigments FCX, PCE and PCC to the discrimination of the classes Chlorophyceae, Cyanophyceae and Bacillariophyceae from their $R_{PHY}(\lambda)$. The strongest correlation between any single input wavelength and the second canonical variate is -0.402 for $R_{PHY}^*(400)$ with $R_{PHY}^*(720)$ having a correlation of -0.351. Again this indicates the importance to the success of the analysis of incorporating the end wavelengths for the canonical discriminant analysis of $R_{PHY}(\lambda)$. The other R_{PHY}^* wavelength which stands out in its correlation with the second canonical variate is 656 nm (with a correlation of -0.365) which represents absorption by CHL*b* which is the marker pigment for the Chlorophytes.

The canonical scores for some of the measured $a_{PHY}(\lambda)$ and $R_{PHY}(\lambda)$ consistently fell some distance away from the class mean (in canonical space) regardless of the wavelength set or data

transformation. In Figure 66 (the analytical $\ln(a_{PHY}(\lambda)^*)$ data set) *Dictiosphaerium* lay closer to the Bacillariophytes than the Chlorophytes, two of the *Fragilaria* spectra fell closer to the Cryptophytes than to the Bacillariophytes, one *Rhodomonas* and one *Cryptomonas* spectra fell closer to the Bacillariophytes than the Cryptophytes, and two *Euglena* spectra fell nearer to the Chlorophyte class mean than to that of the Euglenophytes. In most cases these outliers arise because the canonical discriminant analysis is attempting to separate classes on the basis of pigment absorption when the pigments composition of some of those classes is very similar. Another reason for the poor performance of these particular spectra may be pigmentation changes through photoadaptation (§ 2.1.5.2). However this is unlikely because the culturing conditions were similar for all species. Alternatively cellular senescence may have altered the pigment composition.

Excluding outliers from the canonical discriminant analysis obviously improves the discriminatory capacity of the resulting canonical coefficients. The analytical wavelength set applied to the $\ln(a_{PHY}^*(\lambda))$ resulted in a canonical discrimination with a Wilks' Lambda of 0.002 and first and second canonical correlations of 0.982 and 0.929 (compared with a Wilks' Lambda of 0.004 and canonical correlation of 0.973 and 0.969 when the outliers were included). Furthermore the between-class separation and within-class clustering is improved when outliers are dropped from the input data. It is likely that removing the variance caused by these outliers may achieve better results when the canonical discriminant analysis is used for predicting class.

7.2.4.1 Classifying spectra using canonical discriminant analysis

The canonical coefficients derived from the pure culture $a_{PHY}(\lambda)$ (excluding outliers) using the analytical wavelength set (Table 27) can be used as an identification routine *i.e.* the coefficients developed from known spectra can be used to calculate the canonical scores for unknown spectra. Class assignment can be made on the basis of proximity to the nearest class mean as recorded in canonical space for the spectra used in the known data set.

When this assignment rule was used to identify the class of the pure culture $a_{PHY}(\lambda)$ data set (transformed to $\ln(a_{PHY}^*(\lambda))$ 92% of the 88 spectra were correctly classified. This rises to 96% when only those 74 spectra from the three main classes were considered (Chlorophyceae, Cyanophyceae and Bacillariophyceae). When the same assignment rule was applied to the 50 mixed culture absorption spectra transformed to $\ln(a_{PHY}^*(\lambda))$ 54% of the spectra had their dominant class correctly identified (Figure 70). 60% of the misassignments involved classifying a Euglenophyceae dominated spectra as a Chlorophyceae dominated spectra. Restricting the mixed culture $a_{PHY}(\lambda)$ data set to the 35 spectra composed from the three main phytoplankton classes only improved the identification of the dominant class to 81%. This improvement principally arose from the lack of misassignments between the classes Chlorophyceae and Euglenophyceae.

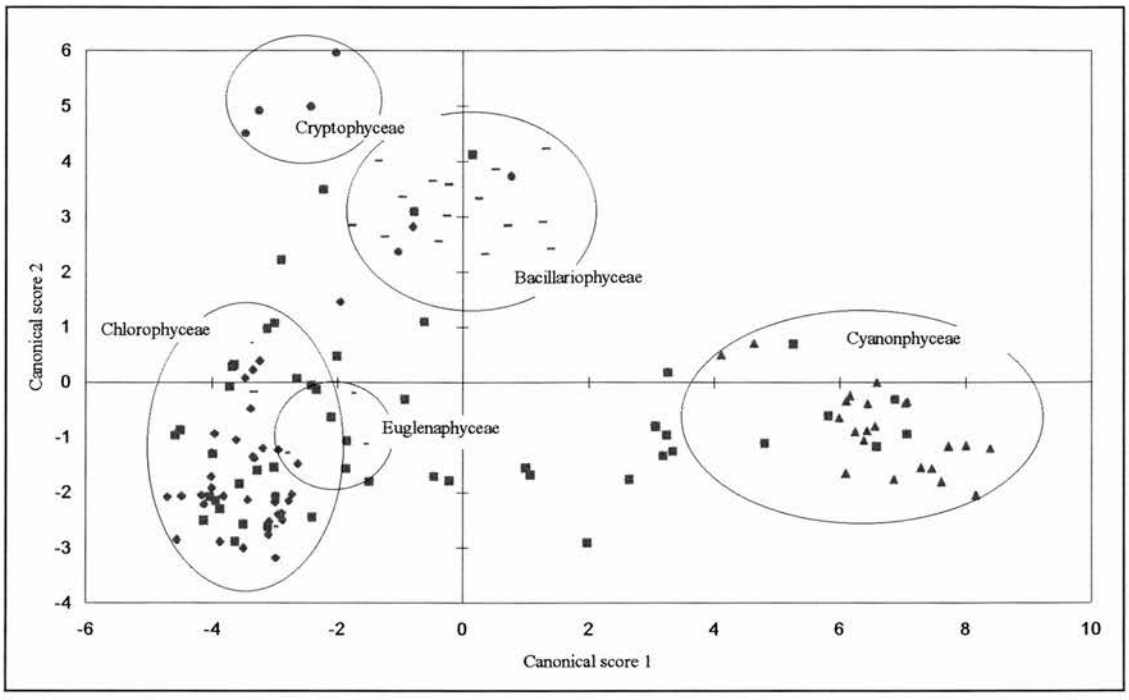

Figure 70. The standardised canonical scores for the 50 mixed culture $a_{PHY}(\lambda)$ (calculated using the canonical coefficients in Table 27) superimposed upon the standardised canonical scores for the pure culture $a_{PHY}^*(\lambda)$ for which the hand drawn ellipses highlight individual class groupings. (Key: diamonds = Chlorophytes, triangles = Cyanophytes, long dashes = Bacillariophytes, circles = Cryptophytes, short dashes = Euglenophytes and large squares = mixed culture spectra).

An alternative classification procedure using canonical discriminant analysis is to use the relative distance of an 'unknown' spectra to class means when plotted in canonical space. Calculation of the relative distance involves ascertaining the total distance from each canonical score to each class mean when plotted in canonical space. The proportion of this total distance (as a percentage) can then be used to assess the relative distance of each variable to each class mean. Hence the relative distance is similar to the likelihood of class membership with smaller values indicating higher probability of the spectra coming from the particular class. Figure 71 shows the application of both the absolute and fuzzy classification techniques for an $a_{PHY}(\lambda)$ from a mixture of Sphaerocystis and Cryptomonas at a ratio of 1 : 2 as established from their CHL*a* concentrations. The canonical scores for this spectra using the analytical wavelength set and the $\ln(a_{PHY}^*(\lambda))$ transformation place the spectra closest to the Cryptophyceae. However the spectra is virtually equidistant from four of the five classes, suggesting that the absolute classification is too precise.

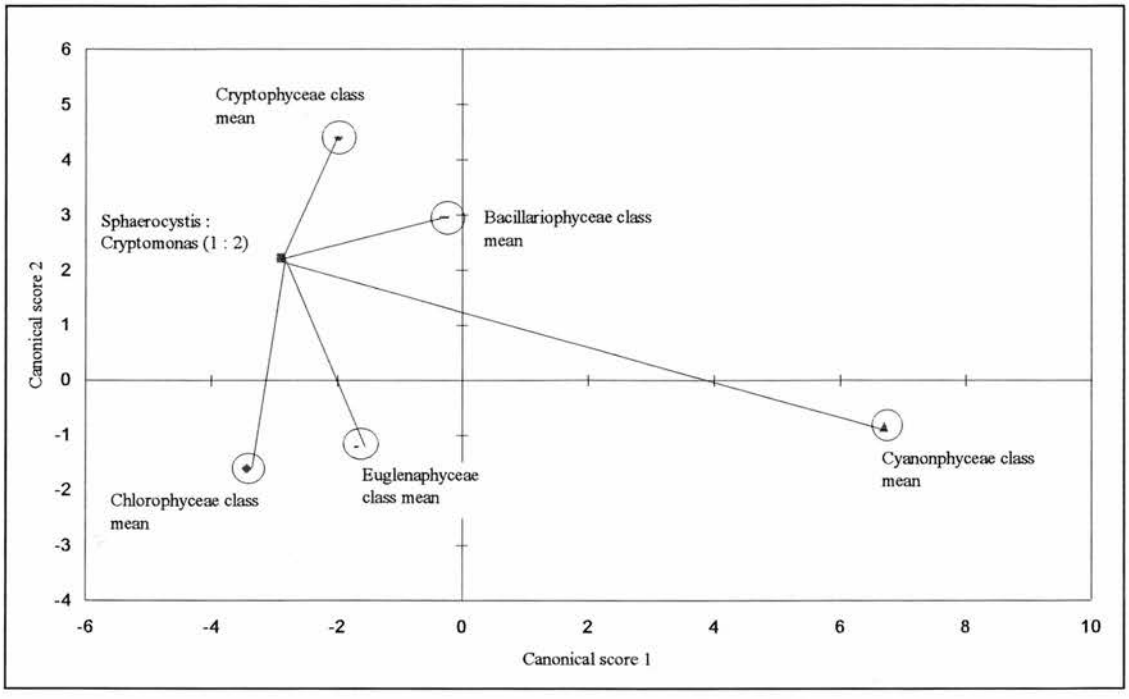

Figure 71. The canonical scores for a mixed culture $a_{PHY}(\lambda)$ (Sphaerocystis : Cryptomonas at 2 : 1) calculated using the canonical coefficients from the analytical wavelength set for the five class $\ln(a_{PHY}^*(\lambda))$ and shown with respect to the class mean canonical scores calculated using the same canonical coefficients.

The fuzzy assignment routine improves if the number of classes is reduced because the canonical discriminant analysis can maximise the distance between-class means in canonical space making misassignments less likely. Using the mixed culture $a_{PHY}(\lambda)$ database restricted to the three main classes and applying the canonical coefficients devised from the three class $\ln(a_{PHY}^*(\lambda))$ data set, the fuzzy assignment system identified the dominant class in 83% (29 out of 35) of spectra. It also correctly identified the sub-dominant class in 43% (15 out of 35) of the spectra.

One drawback to the relative distance classification is the reliance upon the apparently arbitrary location given to each class mean in canonical space by the canonical discriminant analysis. In Figure 66 the Cyanophyte class mean is located some distance from the other classes. Consequently, if the sample spectrum is not dominated by a Cyanophyte then the relative distance to the Cyanophyte class mean will be high and there will be little chance of picking-up the subtle presence of Cyanophytes in that spectrum when using this technique. Alternatively, for a spectra dominated by a Chlorophyte the distance to the Euglenophyte class mean will always be small so the relative distance assignment routine will always be biased towards suggesting a Bacillariophyte rather than a Cyanophyte. It is apparent that for the relative distance classification to be truly effective the class means in canonical space would need to be equidistant. Given the class mean distribution in Figure 66, perhaps the best that can be achieved is an absolute identification of the dominant class with an indication of the

presence of other classes. As with the class identification ratios in the previous section the success of the application of canonical discriminant analysis can be improved if the number of classes is reduced through *a priori* knowledge of the likely presence of particular classes. For instance in a particular lake at a particular time of the year the presence of Euglenophytes may be unlikely, for such lakes the classification would be improved if the canonical discriminant analysis was applied using canonical variates developed in the absence of Euglenophytes.

7.2.5 Summary of Canonical discriminant analysis

Canonical discriminant analysis has been successfully applied to the problem of identifying phytoplankton class from both pure culture $a_{PHY}(\lambda)$ and $R_{PHY}(\lambda)$, and mixed culture $a_{PHY}(\lambda)$. In Chapter Eight the technique is applied to reflectance spectra from a natural water body.

Stepwise discriminant analysis showed the importance of certain wavelengths to the discrimination of phytoplankton class. These wavelengths were those where marker pigments were reported to absorb. Given this reliance on marker pigment absorption wavelengths it was possible to develop a set of wavelengths which proved adept at between-class discrimination. Such a wavelength set should be more universally applicable than the stepwise selected wavelengths when using canonical discriminant analysis as a classification tool because they were not exclusively drawn from the restricted pure culture $a_{PHY}(\lambda)$ data set.

Foody (1996) notes that a significant drawback to the general application of canonical discriminant analysis to the classification of remote sensing data is the reliance that must be placed upon the class distributions as measured in the sample data set. The pure culture $a_{PHY}(\lambda)$ and $R_{PHY}(\lambda)$ data sets are not comprehensive. Increasing both the species and class range of these data sets would improve the development of the canonical coefficients used for classification purposes thereby improving the prediction propensity of this technique.

7.3 Using derivative analysis to classify phytoplankton reflectance spectra

Derivative analysis was first used as a spectral resolution enhancement technique in analytical spectroscopy (O'Haver 1982, Shukla and Rusling 1984). Converting spectra into a derivative accentuates the minor spectral components which are masked in raw spectra by the presence of strong background signals thereby improving the detectability of otherwise difficult to identify features (Martin 1957, Butler and Hopkins 1970, O'Haver and Green 1976, O'Haver 1982).

In the field of aquatic sciences derivatives have been used to highlight *in situ* $a_{PHY}(\lambda)$ features in $a_{TOT}(\lambda)$ and to highlight pigment features in *in vivo* $a_{PHY}(\lambda)$ when those pigment absorption centres are over-lapping (Faust and Norris 1982 and Richardson *at al.* 1994). Derivatives have also been used to assess phytoplankton population structures from the $a_{TOT}(\lambda)$ of natural water bodies through

identification of their more subtle accessory pigment absorption features (Bidigare *at al.* 1989, Richardson *at al.* 1994, Smith and Alberte 1994). O'Haver (1991) used derivatives to obtain a high resolution 'fingerprint' of *in vivo* $a_{PHY}(\lambda)$ or $R_{PHY}(\lambda)$.

Derivatives have also been used in remote sensing applications. Demetriades-Shah *at al.* (1990) used derivatives to analyse the red-edge in vegetation reflectance spectra. The work of Richardson *et al.* (1994), Richardson *et al.* (1995) and Richardson and Ambrosia (1997) describes the use of the derivatives to identify phytoplankton accessory pigments in the coastal environment of Florida Bay, USA. By using a combination of remote sensing instruments at various platforms with ground truth spectroradiometric measurements and HPLC analysis of pigments they were able to relate derivative features to the spectral signatures of class marker pigments and thereby identify dominant class type.

This section describes the application of derivative analysis to the problem of identifying class from phytoplankton absorption and reflectance spectra. A 4^{th} derivative spectra is used to obtain a 'finger print' of $a_{PHY}(\lambda)$ for each class and the presence of class distinctive marker pigment absorption features are identified for the purpose of classification. The technique is then applied to $R_{PHY}(\lambda)$ much in the same manner as Richardson *et al.* (1994).

7.3.1 Properties of derivatives

Derivatives describe the change in slope of the input spectra by wavelength. For raw, untransformed $a(\lambda)$ (the zeroth derivative spectra), the 1^{st} derivative is $\delta a/\delta\lambda$, higher order derivatives are $\delta^n a/\delta\lambda^n$ where n = 2, 3 or 4 for the 2^{nd}, 3^{rd} and 4^{th} derivatives respectively. The 1^{st} derivative describes the slope of a tangent to the zeroth derivative along the length of that spectra. The 2^{nd} derivative is a measure of the curvature or rate of change of the slope of the zeroth spectra. 3^{rd} and 4^{th} derivatives continue in this manner, each successive derivative describing the slope of the n^{th} -1 derivative.

Derivatives enhance subtle features in spectral data because a property of derivatives is that a peak in the n^{th} derivative is directly proportional to the amplitude of that peak, and inversely proportional to the n^{th} power of its width, in the zeroth derivative (O'Haver 1991). It therefore follows that differentiation will discriminate against broad spectral features in favour of narrow-peak components in the zeroth derivative. When dealing with the $a_{PIG}(\lambda)$ components of $a_{PHY}(\lambda)$ the ability of derivatives to improve spectral resolution is dependent upon the breadth of individual pigment component absorption bands (Butler and Hopkins 1970, Smith and Alberte 1994).

The ability of derivatives to enhance high frequency signals at the expense of low frequency signals is advantageous to remote sensing. Algorithms based upon spectral ratios or spectral differences will only correct for background signals if those signals are spectrally on the same magnitudinal level or on a constant slope. Background signals with their own complex spectral form will interfere with

ratio-based algorithms. Derivatives can suppress low frequency background noise and will therefore be unaffected by the influence of background spectral components (Demetriades-Shah *at al.* 1990). For example if there are two signals combined in a spectra Y, with y_1 being a simple background signal and y_2 being a more complex polynomial:

$$y_1 = a_1 + b_1x \quad \text{.....7.3.1.1}$$

$$y_2 = a_2 + b_2x + c_2x^2 + d_2x^3 \quad \text{.....7.3.1.2}$$

and each is at a particular concentration C, then the combined signal Y is:

$$Y = C(y_1) + (1 - C)y_2 \quad \text{.....7.3.1.3}$$

The first derivative Y is $\delta Y/\delta x$:

$$\delta Y/\delta x = C\delta y_1 / \delta x + (1 - C)\delta y_2 / \delta x$$

$$= Cb_1 + (1 - C)(b_2 + 2c_2x + 3d_2x^2) \quad \text{.....7.3.1.4}$$

Thus, on calculation of the 1st derivative the signal y_1 has been reduced to the constant Cb_1. Calculating the second derivative will reduce the contribution of y_1 to the concentration coefficient C which only affects the magnitude of the 2nd derivative having no influence upon its spectral form:

$$\delta^2Y/\delta x^2 = (1 - C)(2c_2 + 6d_2x) \quad \text{.....7.3.1.5}$$

Thus, the effect of increasing the derivative will be to remove the more complex background signals. In terms of aquatic optics if the original input spectra was the visible part of a $a_{TOT}(\lambda)$ from a natural water body consisting of absorption signals from $a_{H2O}(\lambda)$, $a_{DYS}(\lambda)$ and $a_{SES}(\lambda)$, then the 2nd derivative will remove the influence of the exponential $a_{DYS}(\lambda)$ leaving only those contributions from $a_{H2O}(\lambda)$ and $a_{SES}(\lambda)$. This is contrary to the findings of Goodin *at al.* (1993) who mistakenly suggest that the spectral components of the complex $a_{H2O}(\lambda)$ will be removed by calculating the 1st derivative.

The 2nd and 4th derivatives have been most frequently used for the purpose of pigment identification in $a_{PHY}(\lambda)$ and $R_{PHY}(\lambda)$ (Faust and Norris 1982, Owens *at al.* 1987, Richardson *at al.* 1994, Smith and Alberte 1994). The location of 2nd derivative peaks and troughs roughly match the location of features in the zeroth derivative but this is achieved more precisely by the 4th derivative. A 2nd derivative peak represents a trough in the zeroth derivative while a 4th derivative peak represents a peak in the zeroth derivative thereby allowing a more intuitive interpretation of the derivative features. Furthermore the 4th derivative is less affected by overlapping $a_{PIG}(\lambda)$ and achieves a higher degree of spectral resolution (Bidigare *at al.* 1989).

A drawback of using high frequency input signals is that the calculation of derivatives will lead to noise amplification. If the derivative is calculated in its most basic form then with each successive derivative the noise from the original spectra will be doubled (Demetriades-Shah *at al.* 1990, O'Haver 1991). However, the amplification of noise can be mitigated by using more advanced derivative calculation routines which include smoothing routines (O'Haver 1982).

The ability of derivatives to separate two overlapping $a_{PIG}(\lambda)$ bands requires consideration. This primarily depends upon the properties of the both peaks. If the intensity of the secondary peak is minor and is separated from the primary peak by less than the band-width of the primary peak, then the secondary peak may only appear as a shoulder in the derivative spectra. Peaks separated by only 40% of their half-widths are unlikely to be resolved by the 4th derivatives. Furthermore, the intensities of the derivative peaks generated by overlapping $a_{PIG}(\lambda)$ bands are not related to the intensities of the original peaks of the zeroth derivative. Butler and Hopkins (1970) showed that the spurious derivative peaks could be generated if two absorption bands, located with the same maximum wavelength but with different half-widths are superimposed. Consequently, the validity of peaks in the derivative spectra requires confirmation in both the zeroth and other derivative spectra. Spurious peaks can be generated in the 4th derivative if the zeroth derivative is composed of a series of Gaussian curves. The properties of the Gaussian curves may lead to 'ripple' effects on either side of the main derivative peak. These ripples may then create pseudo peaks which can be misinterpreted as separate features in the zeroth derivative. This is illustrated in Figure 72 which shows a synthesised $a_{PHY}(\lambda)$ composed of two Gaussian absorption curves centred on 440 and 675 nm (representing the two main centres for CHL*a* absorption). The 4th derivative feature generated by the Soret absorption band is a shallow wave, the 4th derivative feature generated by the far-red absorption band is a much sharper peak with surrounding 'ripples' which are significant enough to be misinterpreted as peaks in their own respect.

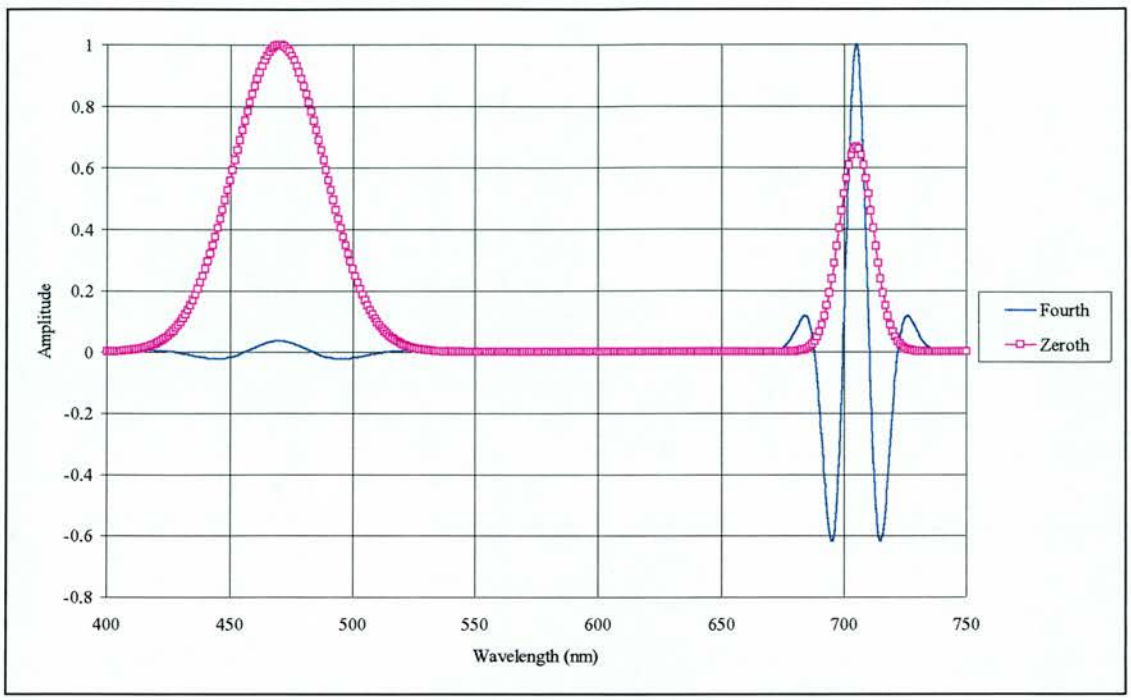

Figure 72. The 4th derivative of a hypothetical phytoplankton absorption spectra (the zeroth derivative) absorbing at the two main CHL*a* absorption centres only. Both spectra have been normalised to unity.

7.3.2 Methods including the Savitzky-Golay derivative calculations

The 4th derivatives in this study were calculated according to Savitzky and Golay (1964) as modified by Turrell (1982). The Savitzky-Golay method is the most frequently employed form of derivative calculation (Demetriades-Shah *at al.* 1990, Smith and Alberte 1994). It is essentially a polynomial smoothing function adapted for differentiation. The least squares polynomial fitting procedure requires the solving of the line which best fits the data and can be represented by the following equation:

$$y = a_0 + a_1x + a_2x^2 + a_3x^3 .... + a_nx^n \quad \text{.....7.3.2.2}$$

Savitzky and Golay (1964) found that generation of the 'a' coefficients could be simply evaluated using a set of integers as a weighting function and the equation could then be used to solve for the central point of the polynomial. This integer method is exactly the same as solving for the full polynomial. However, its simplification meant that the technique could be easily transferred to a computer. Least squares polynomial fits can be used for both smoothing and the generation of derivatives. In their paper Savitzky and Golay published the integers necessary for least squares fitting which are dependent upon the order and width of the polynomial. Turrell (1982) simplified the derivative calculation by publishing a method for computing these integers in matrix form. A Turbo Pascal v7.0 program (*Savitzky.pas*) written by T.J.Malthus used to facilitate both the least squares smoothing of data and the calculation of various derivatives. *Savitzky.pas* allowed the user to choose the parameters required for the least square fitting according to Savitzky and Golay (1964).

The main consideration when using the Savitzky-Golay routine is the choice of parameters necessary for the Savitzky-Golay calculation. This is the decision of the operator and is more often than not a subjective process (Enke and Nieman 1976). The following section discusses these parameters which are polynomial width, polynomial order and required derivative. The user can also specify whether the required order of derivative was to be calculated directly or reached by successive passes of a first order derivative to achieve the same result.

The choice of polynomial width determines both the degree of smoothing and the number of data points that are lost from the end of each spectrum. The width of the polynomial dictates the number of points that are lost at the ends of the input spectra, for a 2n + 1 point Savitzky and Golay smooth there are n points lost from each end of the spectra. The importance of the polynomial width parameter is that subtle features in the raw spectra are lost through smoothing as the width increases. Furthermore as width increases data loss from each end of the spectrum increases proportionally causing end point distortions to encroach at either end of the spectra (Enke and Nieman 1976, Khan 1987). It should also be noted that the magnitude of a derivative feature is a function of the polynomial width; increasing width causes a reduction in derivative peak height and trough depth. This reductions is particularly evident for the higher order derivatives.

Choosing a higher order polynomial increases the resolution of subtle features in the raw spectra. It should also be noted that there is no difference between derivatives calculated using either the 2^{nd} or 3^{rd} order polynomial, or between derivatives calculated using the 4^{th} and 5^{th} order polynomials. This concurs with the tables of coefficients published by Savitzky-Golay (1964) where quadratic and cubic polynomials have the same coefficients, as do the quartic and quintic order polynomials.

Assuming that the user knows which derivative is required, then there is a consideration of whether to calculate the higher order n^{th} derivatives directly or whether to calculate them cumulatively by passing a 1^{st} order derivative through the data n times. Trials involving the calculation of a 4^{th} derivative $a(\lambda)$ with four passes of a 1^{st} derivative showed that severe distortions encroach from either end of the spectra and these distortions increase with polynomial width.

To assess the ability of 2^{nd} and 4^{th} derivatives to separate compound pigment features a synthesised $a_{PHY}(\lambda)$ was constructed by summing seven variously shaped Gaussian bands as shown in Figure 73. Judicious placement and summation of these Gaussian bands allows the construction of a curve of similar shape to a generalised $a_{PHY}(\lambda)$. The 2^{nd} and 4^{th} derivative of a synthesised phytoplankton absorption curve was calculated using a 15 point 4^{th} order polynomial. The location of the troughs in the 2^{nd} derivative and the peaks of the 4^{th} derivative shown in Figure 73 were compared with the location of the centre points of the Gaussian curves. These peak and trough locations are summarised in Table 30. The 2^{nd} derivative identified the centres of the Gaussian curves more precisely than the

4th derivative but failed entirely to identify the 655 nm centred band, although this feature was also imprecisely located by the 4th derivative. Figure 73 shows this 655 nm band to be a very broad feature and such features are not well separated by derivative analysis. Similarly the broad 490 and 570 nm bands were imprecisely located by both derivatives. The sharper, but magnitudinally weaker, 530 nm band was both prominently visible and accurately identified in both derivatives. The most prominent feature in both derivatives was the magnitudinally strong but narrow Gaussian band at 676 nm.

Table 30. The locations of the centres of the Gaussian bands and the corresponding 2nd and 4th derivative features shown in Figure 73.

Gaussian band	*Location of centre point of Gaussian band (nm)*	*Location of 2nd derivative trough (nm)*	*Location of 4th derivative peak (nm)*
1	440	440	439
2	490	494	498
3	530	530	530
4	570	571	565
5	625	625	624
6	655	not present	644
7	676	676	676

Given that the 2nd derivative may miss some absorption features and, in accordance with the work of Owens *at al.* (1987), Bidigare *at al.* (1989), Richardson *at al.* (1994) and Owens and Alberte (1994) all of whom used the 4th derivative, the 4th derivative will be used for this study. However, the imprecision with which the 4th derivative identifies peak locations must be borne in mind for the rest of this study.

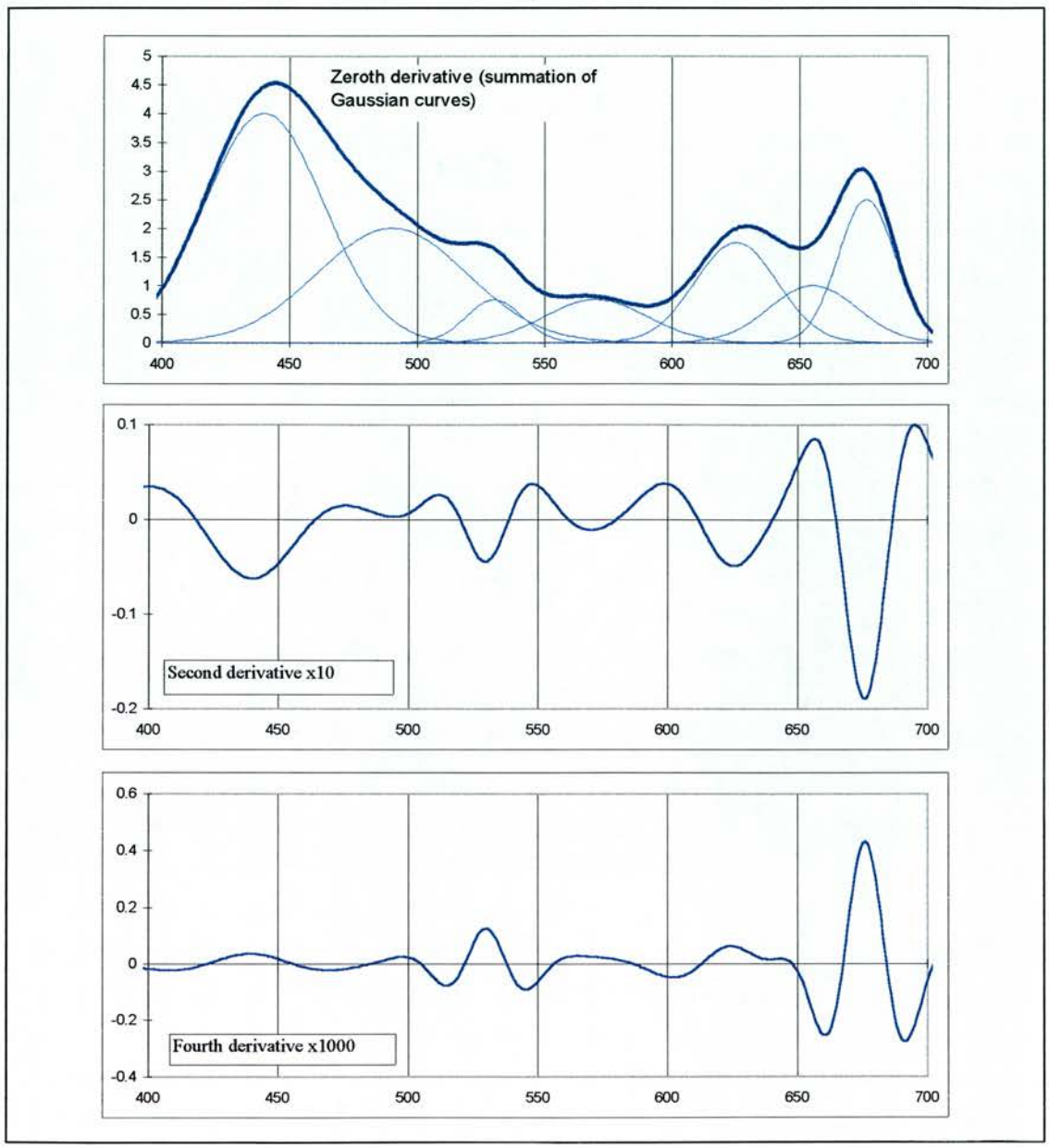

Figure 73. 2nd and 4th derivatives of a synthesised absorption spectrum comprising of a summation of seven Gaussian bands centred on the wavelengths given in Table 30. The *y*-axis units are arbitrary and the magnitude of the derivatives have been augmented for display purposes. Note the coincidence of the 2nd derivative troughs and 4th derivative peaks with the centres of the Gaussian bands.

The 4th derivatives were all calculated using a 4th order polynomial with a 15 or more point smooth depending upon the amount of noise in the data set. Henceforth a Savitzky-Golay 4th derivative calculated using a 4th order 29 point wide polynomial will be denoted according to the following standard {4,4,29,1} corresponding to {derivative, polynomial, width and number of passes}.

7.3.3.1 Using derivative analysis to classify phytoplankton absorption spectra

4th derivatives were calculated for the pure and mixed culture $a_{PHY}(\lambda)$ and the location of derivative peaks identified. Given that $a_{PHY}(\lambda)$ is a combination of the absorption curves of the individual pigments and each peak in the 4th derivative represents one of these absorption centres, then location of 4th derivative peaks should allow identification of pigments. This was achieved by cross-referencing the 4th derivative peak locations against the wavelength locations of the peaks in *in vivo* pigment absorption spectra, as collated in the pigment absorption database (Appendix Two). Identification of pigment composition and in particular identification of the class distinctive marker pigments should lead to successful classification of the spectra. Two fully worked examples of this procedure are given below, one for the $a_{PHY}(\lambda)$ of the pure *Anabaena* culture measured in the experimental tank, and one for the $a_{PHY}(\lambda)$ of a mixture of two cultures.

Figure 74 shows the raw and 4th derivative {4,4,29,1} *Anabaena* $a_{PHY}(\lambda)$ spectra. The pigments responsible for the peaks in the 4th derivative spectra have been identified through the pigment database. Bearing in mind the imprecision with which 4th derivatives identify features in the raw spectrum, the 4th derivative peak locations were identified at a 5 nm resolution. Peak locations were cross-referenced against the absorption properties for the pigments in the pigment database (Table 31). In some cases two pigments are reported as absorbing at the same wavelength, as was the case for the 582 nm peak which has been identified as either PCC or PCE. The combination of identified pigments was compared to the reported pigment compositions (Table 1 and Appendix One) and the class assigned accordingly. The combination of pigments identified in the *Anabaena* $a_{PHY}(\lambda)$ were indicative of a Cyanophyte which are dominated by the biliproteins, in particular the Cyanophyte marker pigment PCC at 640 nm. The presence of the FCX in a Cyanophyte has not previously been reported in the literature, however FCX is the only pigment reported as absorbing around 530 nm. It is more likely that the peak was an unreported centre of absorption for MYX which is a Cyanophyte marker pigment.

Table 31. The locations of 4th derivative peaks and pigments associated with those peaks for the *Anabaena* $a_{PHY}(\lambda)$ as shown in Figure 74.

4th derivative peak location (nm)	*Pigments reported as absorbing at this wavelength*
440	CHL*a*
485	β-CART, various xanthinols
525	FCX, various xanthinols
560	PCE, PCC
580	PCE, PCC
615	PCC, APC
640	PCC, APC
680	CHL*a*

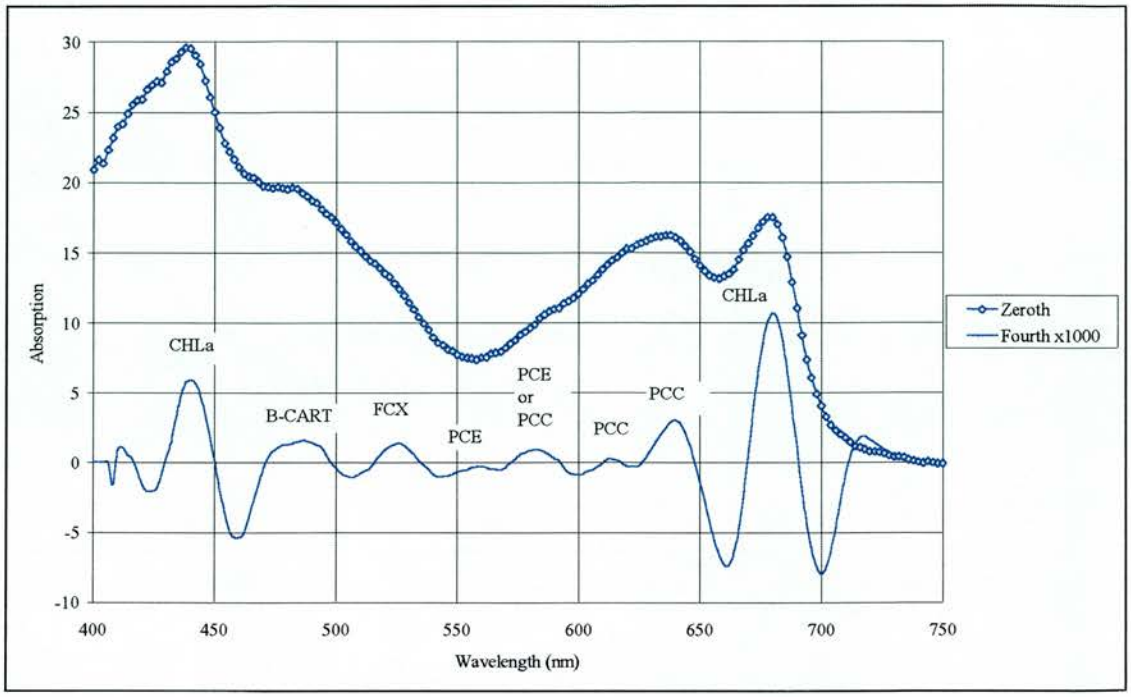

Figure 74. The zeroth and 4th derivative {4,4,29,1} *Anabaena* $a_{PHY}(\lambda)$. The 4th derivative spectra has been annotated to show the pigment most likely to be responsible for that absorption feature, these are summarised in Table 31.

Figure 75 shows the 4th derivative {4,4,29,1} for the $a_{PHY}(\lambda)$ of a mixture of *Anabaena* and *Selenastrum* (55% : 45%). Again the pigments most likely to be responsible for the 4th derivative peaks have been annotated. These peaks and their associated pigments are summarised in Table 32. The Cyanophyte biliproteins dominate the $a_{PHY}(\lambda)$ while the Chlorophyte marker pigment CHL*b* could not be separated from the biliproteins even at 650 nm where it absorbs almost uniquely. The width of the polynomial had to be reduced to 15 points before a 650 nm CHL*b* peak became visible. However at this relatively narrow smoothing width signal noise creates spurious peaks. Derivative analysis of the pure *Selenastrum* $a_{PHY}(\lambda)$ that was used in the mixture also failed to identify the CHL*b* 650 nm absorption centre at the broader polynomial widths.

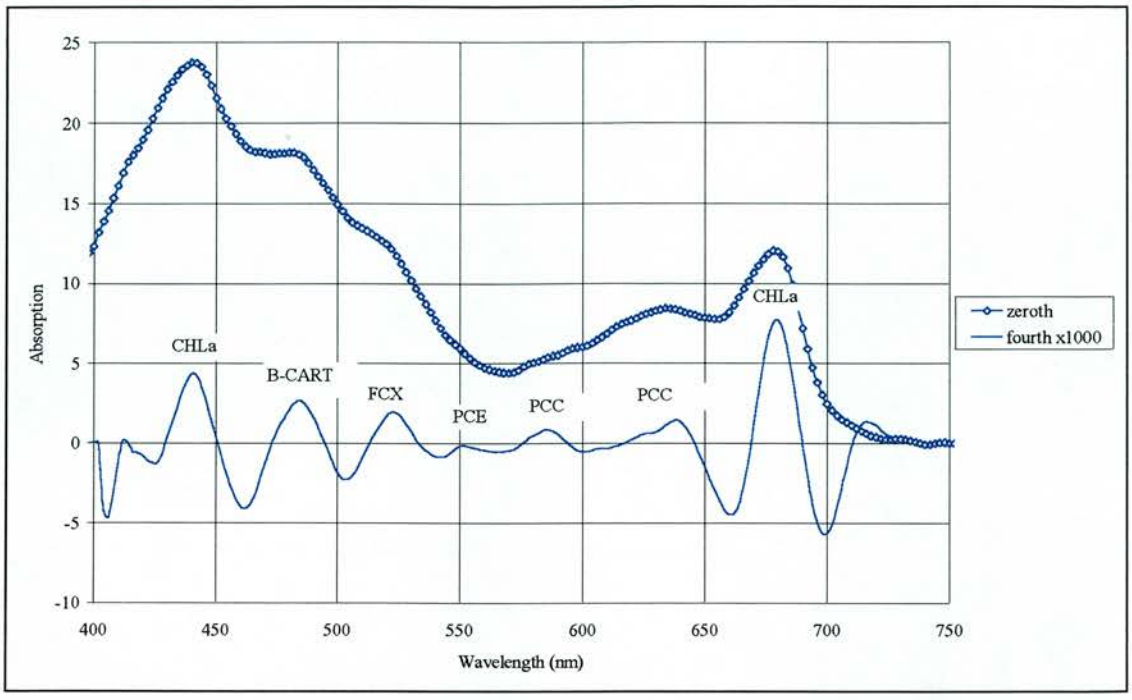

Figure 75. The zeroth and 4th derivative {4,4,29,1} for an $a_{PHY}(\lambda)$ mixture of *Selenastrum* and *Anabaena*. The pigments most likely to be responsible for the 4th derivative peaks have been annotated and are summarised in Table 32.

Table 32. The 4th derivative {4,4,29,1} peaks and associated pigments for the $a_{PHY}(\lambda)$ for a mixture of *Anabaena* and *Selenastrum* (55% : 45%) as shown in Figure 75.

4th derivative peak location (nm)	*Pigments reported as absorbing at this wavelength*
440	CHL*a*
485	β-CART, CHL*b*
520	FCX, various xanthinols
550	PCE, PCC
585	PCC
640	PCC, CHL*b*, APC
680	CHL*a*

7.3.3.2 Using derivative analysis to classify phytoplankton reflectance spectra

The 4th derivative {4,4,29,1} was calculated for the $R_{PHY}(\lambda)$ measured for the same *Anabaena* culture used in Figure 74. For comparison purposes both the *Anabaena* $R_{PHY}(\lambda)$ and $a_{PHY}(\lambda)$ 4th derivatives are shown in Figure 76. The peaks in the 4th derivative $a_{PHY}(\lambda)$ have already been identified as centres for absorption for pigments, because $R_{PHY}(\lambda)$ is principally the inverse of $a_{PHY}(\lambda)$, the pigment absorption features in the $R_{PHY}(\lambda)$ are transposed into 4th derivative troughs. The 4th derivative reflectance troughs were identified at a 5 nm resolution and their locations cross-referenced to reported pigment absorption properties. The 4th derivative $R_{PHY}(\lambda)$ troughs in Figure 76 have been annotated with the associated pigment and are summarised in Table 33.

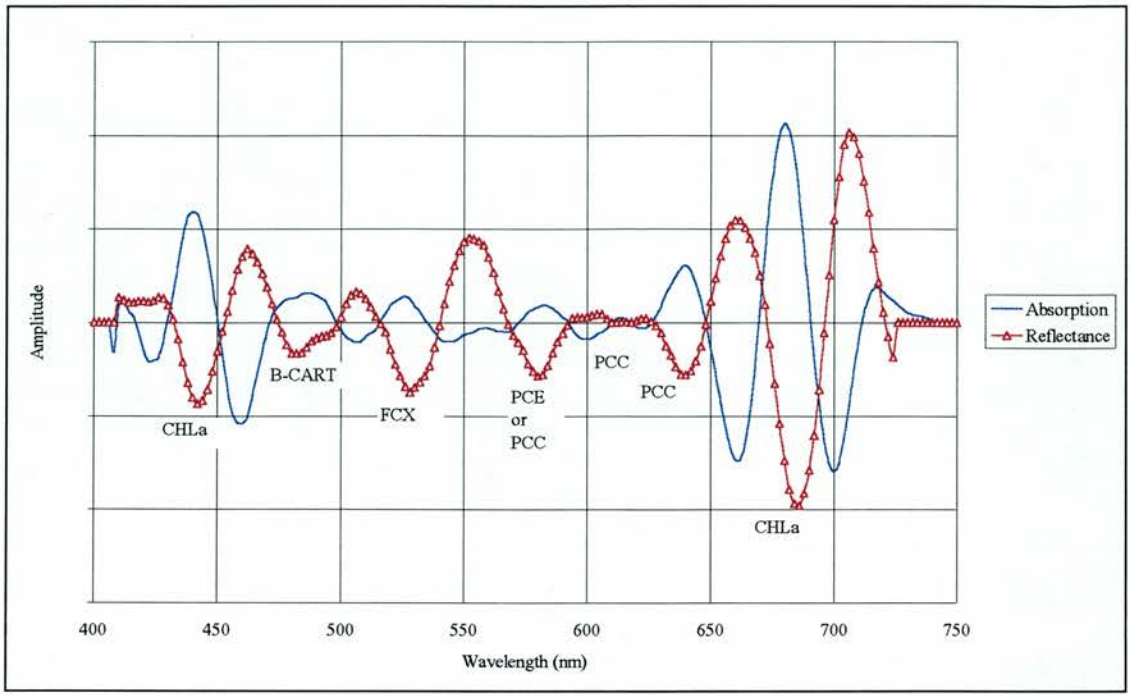

Figure 76. The 4th derivative for *Anabaena* $a_{PHY}(\lambda)$ {4,4,29,1} and $R_{PHY}(\lambda)$ {4,4,29,1}. The $a_{PHY}(\lambda)$ 4th derivative was previously shown in Figure 74. The pigments responsible for the troughs in the $R_{PHY}(\lambda)$ 4th derivative have been annotated. Note that the spectra have been magnitudinally altered for display purposes.

Table 33. Comparison of the location of the 4th derivative {4,4,29,1} $a_{PHY}(\lambda)$ peaks and 4th derivative $R_{PHY}(\lambda)$ troughs for the *Anabaena* culture used in the experimental tank (Figure 74 and Figure 76 respectively). The pigments with absorption centres located near these features are given.

Location of absorption peaks (nm)	*Pigment associated with absorption feature*	*Location of reflectance troughs (nm)*	*Pigment associated with reflectance feature*
440	CHL*a*	440	CHL*a*
485	β-CART, various xanthinols	480	β-CART, various xanthinols
525	FCX	525	FCX
560	PCE	580	PCE
580	PCE	595	PCE
615	PCC	615	PCC
640	PCC	640	PCC
680	CHL*a*	685	CHL*a*

With the exception of the 560 nm feature all the 4th derivative $a_{PHY}(\lambda)$ peaks identified in Table 31 were present as $R_{PHY}(\lambda)$ troughs. There was a slight relocation of some features as shown in Table 33. A 4th derivative $R_{PHY}(\lambda)$ feature that was not present in the 4th derivative $a_{PHY}(\lambda)$ was identified at 595 nm.

7.3.4 Results of the derivative analysis

4th derivatives {4,4,29,1} were calculated for the 88 pure culture phytoplankton absorption spectra. The peaks in each spectrum were identified to a resolution of 5 nm. The location of the 4th derivative

peak features are shown for each class in Table 34 to Table 38. The percentage of the within-class spectra exhibiting that feature and for comparison purposes, the percentage of within-class spectra from the other classes exhibiting a 4th derivative peak at that wavelength location are also given along with the pigments reported as absorbing at those wavelengths.

All the classes have 4th derivative peaks at 440 nm confirming the ubiquitous presence of CHL*a* which absorbs both at this wavelength and in the far-red. However, the location of the far-red 4th derivative peak was class dependent, occurring at 675 nm in the Bacillariophytes and Cryptophytes and at 680 nm in the Chlorophytes, Cyanophytes and Euglenophytes. The shorter wavelength location for the far red CHL*a* peak in the Bacillariophytes and Cryptophytes adds further weight to the hypothesis of Wróbel (1977) who found that on mixing a pure CHL*a* extract with a pure CHL*c* extract in the absence of any solvent (so avoiding hypsochromatic and bathychromatic shifts in peak pigment absorption) the far-red absorption peak shifted from > 680 to 676 nm suggesting that some aggregation between these pigments occurs which slightly alters the spectral properties of the former. Both these classes are reported to contain CHL*c*. The other 4th derivative peak feature present in all the classes is at 540 nm, consistently present in the Euglenophyceae and Chlorophyceae but less consistently present in the Bacillariophyceae and Cryptophyceae and infrequently present in the Cyanophyte spectra. This wavelength is the absorption centre for xanthinols and biliproteins, both of which are reported as being present in most classes of freshwater phytoplankton. A 540 nm 4th derivative feature was consistently present in the $a_{PHY}(\lambda)$ measured by Smith and Alberte (1994) who associated it with the absorption by FCX.

When considered within each class the 4th derivative features were seldom consistently present in each sample species spectra with the exception of the $a_{PHY}(\lambda)$ from the class Euglenophyceae. For example, in Table 35 the 485 nm 4th derivative peak was only present in 52% of the Cyanophyte $a_{PHY}(\lambda)$.The consistency with which a feature is exhibited by sample $a_{PHY}(\lambda)$ from each class is important when considering which features should be considered indicative of a class for the purposes of classification.

Of the ten 4th derivative peaks present in the 34 Chlorophytes $a_{PHY}(\lambda)$ (Table 34) those at 490, 510 and 650 nm were not present in the other classes. There was a distinct similarity between the location of the Chlorophyte peaks and the location of the Euglenophyte peaks. This was to be expected because of the similarity in pigment content between these two classes. The 510 nm peak was inconsistently present in the Chlorophyte $a_{PHY}(\lambda)$ so could not be used as a marker feature. However, the 490 and 650 nm peaks were more frequently present in the Chlorophyte species measured. The 490 nm peak suggests that the presence of various xanthinols, which are reported as being present in the Chlorophytes, may be the ideal class marker feature. The 650 nm peak indicates the presence of CHL*b* which has a double peak in *in vivo* absorption at 470 and 650 nm. The 470 nm peak coincides with the absorption by other pigments and thus would be unsuitable as a marker feature. The 650 nm

CHL*b* absorption peak is in close proximity to the strong far-red CHL*a* absorption and is apparent in the 4th derivative as a shoulder rather than a distinctive peak, contrary to the findings of Faust and Norris (1982), Bidigare *at al.* (1989) and Smith and Alberte (1994) all of whom were able to fully resolve this 650 nm CHL*b* peak in 4th derivative Chlorophyte $a_{PHY}(\lambda)$. This-not-withstanding there are no 650 nm features in any of the other classes, particularly the Euglenophytes which are also reported to contain CHL*b*, consequently this 650 nm feature has the potential to be used as a marker feature for the class Chlorophyceae.

Table 34. 4th derivative peak locations for the 34 Chlorophyte $a_{PHY}(\lambda)$ and the percentages of spectra within each class exhibiting that feature. The pigments reported as absorbing at each wavelength are given. Potential class marker features are highlighted.

Peak wavelength (nm)	*Chloro-phyte*	*Cyano-phyte*	*Bacillario-phyte*	*Crypto-phyte*	*Eugleno-phyte*	*Pigments reported at this wavelength (**bold** indicates an affiliation with class under consideration)*
440	97%	96%	94%	75%	100%	**CHL*a***
470	79%	89%			83%	**CHL*b***, CHL*c*, B-CART, PRD
490	88%					**DDX**, **VLX**, **LTX**, **TRX**, CART
510	35%					PCE
540	97%	26%	76%	75%	100%	**SPX**, PCE, PRD
580	100%				100%	PCE
620	94%				100%	PCC, APC, CHLc
640	29%	48%				**CHL*b***, CHL*c*
650	71%					**CHL*b***, APC
680	100%	100%			100%	**CHL*a***

There were thirteen 4th derivative peak features present in the Cyanophyceae (Table 35) three of which were present in less than 50% of the 27 $a_{PHY}(\lambda)$ measured in this class. For six of the remaining features the coincidence with peaks in other classes precluded their use as definitive Cyanophyte marker features. This leaves four wavelengths where the Cyanophyte 4th derivative peaks existed in the absence of the peaks of other classes: 485, 520, 555 and 630 nm. The 485 nm peak was only present in 52% of the Cyanophyte spectra measured and is thus too inconsistent to be a marker feature. The 520 nm peak is more consistently present however the only pigment reported as absorbing in the vicinity of 520 nm is FCX which is unlikely to be present in the Cyanophytes. Moreover FCX is reported as being present in the Bacillariophytes and Cryptophytes. Consequently the 520 nm peak would also appear to be unsuitable as a class marker feature. The 555 and 630 nm 4th derivative peaks are much more suitable as they suggest the presence of biliproteins which are reported to occur in high quantities in the Cyanophytes. Furthermore peaks at these wavelengths are unique to the Cyanophytes. The 555 nm PCE absorption feature in particular has no other pigments absorbing in close proximity, making it an ideal marker feature.

Table 35. 4th derivative peak locations for the 27 Cyanophyte $a_{PHY}(\lambda)$ and the percentages of spectra within each class exhibiting that feature. The pigments reported as absorbing at each wavelength are given. Potential class marker features are highlighted.

Peak wavelength (nm)	*Cyano-phyte*	*Chloro-phyte*	*Bacillario-phyte*	*Crypto-phyte*	*Eugleno-phyte*	*Pigments reported at this wavelength (**bold** indicates an affiliation with class under consideration)*
440	96%	97%	94%	75%	100%	**CHL*a***
470	89%	79%			83%	**B-CART**, CHL*b*, CHL*c*, PRD
485	52%					**B-CART**, **TRX**, CHL*b*
495	89%		100%		100%	**PCE**, PUB
520	89%		18%	38%		FCX
540	26%	97%	76%	75%	100%	**PCE**, SPX, PRD
555	85%					**PCE**, **PCC**
570	33%		35%			**PCE**
585	85%		100%	50%		**PCC**
610	63%		71%			
630	70%					**PCC**, CHL*c*, PCB
640	48%	29%				**PCC**, CHL*c*, CHL*b*
680	100%	100%			100%	**CHL*a***

There were ten 4th derivative peaks identifiable in the Bacillariophyte $a_{PHY}(\lambda)$ (Table 36). There was a strong similarity between the Bacillariophyte peak locations and the Cryptophyte peak locations confirming the similarity in pigment composition between these classes. None of the peaks occurred in complete isolation from the peak features of the other classes, suggesting that there is no definitive Bacillariophyte marker feature. However, it is possible to use the locations of more than one peak to develop a 'finger-print' for the Bacillariophyceae. For example, the 465 nm Bacillariophyte peak is matched only by a Cryptophyte peak and the 610 nm Bacillariophyte peak is matched by both the Cyanophytes and Euglenophytes but not with the Cryptophytes. The combined presence of these features should therefore indicate the presence of Bacillariophytes.

Table 36. 4th derivative peak locations for the 15 Bacillariophyte $a_{PHY}(\lambda)$ and the percentages of spectra within each class exhibiting that feature. The pigments reported as absorbing at each wavelength are given. Potential class marker features are highlighted.

Peak wavelength (nm)	*Bacillario-phyte*	*Chloro-phyte*	*Cyano-phyte*	*Crypto-phyte*	*Eugleno-phyte*	*Pigments reported at this wavelength (**bold** indicates an affiliation with class under consideration)*
440	94%	97%	96%	75%	100%	**CHL*a***
465	94%			100%		**CHL*c*, B-CART**
495	100%		89%		100%	**PCE**, PUB
520	18%		89%	38%		**FCX**
540	76%	97%	26%	75%	100%	**PCE**, SPX, PRD
570	35%		33%			**PCE**
585	100%		85%	50%		**PCC**
610	71%		63%			**PCC**
635	94%			88%		**CHL*c***
675	100%			100%		**CHL*a***

The conclusion reached for the Bacillariophyte 4th derivatives suggest that although the 500 and 615 nm features are definitive class marker pigments for the Cryptophytes (Table 37), slight inaccuracies

in the measurement of $a_{PHY}(\lambda)$ or idiosyncrasies in the calculation of the 4th derivative may lead to confusion between the Cryptophytes and the Bacillariophytes. Given this caveat these class-unique features, combined with the 565 nm peak feature, remain the most likely marker features for the Cryptophytes. Furthermore, the biliproteins reported as absorbing at these wavelengths are also reported as occurring in the class Cryptophyceae. The consistency with which these particular peak features were found in the pure culture $a_{PHY}(\lambda)$ suggest, that assignment to the class Cryptophyceae should be reasonably accurate and their absence should lead to a more positive assignment of the class Bacillariophyceae.

Table 37. 4th derivative peak locations for the seven Cryptophyte $a_{PHY}(\lambda)$ and the percentages of spectra within each class exhibiting that feature. The pigments reported as absorbing at each wavelength are given. Potential class marker features are highlighted.

Peak wavelength (nm)	*Crypto-phyte*	*Chloro-phyte*	*Cyano-phyte*	*Bacillario-phyte*	*Eugleno-phyte*	*Pigments reported at this wavelength (**bold** indicates an affiliation with class under consideration)*
440	75%	97%	96%	94%	100%	**CHL*a***
465	100%			94%		**CHL*c*, B-CART**
500	100%					**FCX**, **PCE**, MYX
520	38%		89%	18%		**FCX**
540	75%	97%	26%	76%	100%	**PCE**, SPX, PRD
565	100%					**PCE**, **PCC**
585	50%		85%	100%		**PCC**
615	88%					**PCC**
635	88%			94%		**CHL*c***
675	100%			100%		**CHL*a***

Although the peaks present in the Euglenophyte 4th derivative $a_{PHY}(\lambda)$ are consistently present in all the measured Euglenophyte $a_{PHY}(\lambda)$ none are unique to the class (Table 38). Thus none could be used as definitive class marker features and it is again necessary to use the presence of a combination of features to rule out the possibility of spectra belonging to one of the other classes. This could be achieved by combining the presence of the 495 nm peak with the 580 or 620 nm peak. None of the pigments reported as absorbing at these feature location wavelengths are reported as occurring in the Euglenophyceae. However, the consistency with which these features occur suggests that they are real absorption features and, given their location, they are most likely to be some form of previously unreported biliprotein.

Table 38. 4th derivative peak locations for the five Euglenophyte $a_{PHY}(\lambda)$ and the percentages of spectra within each class exhibiting that feature. The pigments reported as absorbing at each wavelength are given. Potential class marker features are highlighted.

Peak wavelength (nm)	*Eugleno-phyte*	*Chloro-phyte*	*Cyano-phyte*	*Bacillario-phyte*	*Crypto-phyte*	*Pigments reported at this wavelength (**bold** indicates an affiliation with class under consideration)*
440	100%	97%	96%	94%	75%	**CHL*a***
470	83%	79%	89%			**CHL*b***, CHL*c*, B-CART, PRD
495	100%		89%	100%		PCE, PUB
540	100%	97%	26%	76%	75%	SPX, PCE, PRD
580	100%	100%				PCE
620	100%	94%				PCC, APC, CHL*c*
680	100%	100%	100%			**CHL*a***

In summary, pure culture $a_{PHY}(\lambda)$ 4th derivative spectra can be assigned to a class using the class marker features listed in Table 39. It is more convenient to reduce the number of class marker features to those in Table 39 than trying to search derivative spectra for the comprehensive presence of derivative features generated by all the pigments listed as being present in each class. For the Chlorophytes and the Cyanophytes the presence of peaks at either of the locations listed in Table 39 would be sufficient. However, for a more accurate class assignment the presence of all the features would be necessary to confirm the class of an unknown $a_{PHY}(\lambda)$.

Table 39. The class distinctive marker features derived from the pure culture 4th derivative $a_{PHY}(\lambda)$.

Class	*4th derivative peak feature locations (nm)*	*Associated pigments*
Chlorophyte	490, 650	xanthinols, CHL*b*
Cyanophyte	555, 630	PCE, PCC
Bacillariophyte	465, 610	CHL*c*, PCC
Cryptophyte	500, 565, 615	PCE, PCC
Euglenophyte	495, 580, 620	unreported biliproteins

7.3.4.1 Using derivative analysis to classify the mixed culture phytoplankton absorption spectra

The 4th derivative {4,4,29,1} was calculated for the mixed culture $a_{PHY}(\lambda)$ described in § 4.2.5. The dominant and sub-dominant classes were assigned using the marker features presented in Table 39. A class was assigned as dominant if the all the features listed in Table 39 were present and as sub-dominant if only some of the listed features were present.

Of the 50 mixed culture $a_{PHY}(\lambda)$ only 24% had both their dominant and sub-dominant class correctly identified. However, 60% had their dominant classes and 42% their sub-dominant classes correctly identified. For 90% of the spectra the presence of one of the two classes was correctly identified. 50% of the spectra dominated by Euglenophytes, 64% of those dominated by Chlorophytes and 73% of those dominated by Cyanophytes were correctly identified. 25% of the spectra dominated by Euglenophyceae were misclassified as Chlorophyceae and 15% of those dominated by Chlorophyceae were misclassified as Euglenophyceae. Thus the similarity in the pigment compositions of these classes cause some, but not excessive, misclassification.

7.3.4.2 Using derivative analysis to classify the pure culture phytoplankton reflectance spectra

4th derivative {4,4,29,1} of the pure culture $R_{PHY}(\lambda)$ (§ 4.3.4) were calculated. The derivative troughs were identified to a resolution of 5 nm and are presented in the tables below. Table 40 shows that eight of the feature locations identified in the 4th derivative Chlorophyte $a_{PHY}(\lambda)$ (Table 34) were present in the 4th derivative $R_{PHY}(\lambda)$ including the class marker features at 490 and 650 nm. However, the 490 nm feature was not consistently present in all the Chlorophyte $R_{PHY}(\lambda)$. Eleven of the features listed in Table 41 were also present in the Cyanophyte $a_{PHY}(\lambda)$ 4th derivative spectra (Table 35) including the two Cyanophyte marker features. Nine of the features were present in the Bacillariophyte $a_{PHY}(\lambda)$ 4th derivative spectra (Table 36) but only the 465 nm Bacillariophyte marker feature identified in the $a_{PHY}(\lambda)$ was also present in the $R_{PHY}(\lambda)$. The strong within-class coincidence of the location of 4th derivative $a_{PHY}(\lambda)$ and $R_{PHY}(\lambda)$ features is encouraging. Scattering trends which do influence $R_{PHY}(\lambda)$ have been removed during the calculation of the 4th derivative.

Table 40. The location of the 4th derivative troughs for the Chlorophyte $R_{PHY}(\lambda)$ and the percentage of the within-class spectra exhibiting that feature. Highlighted wavelengths correspond to the 4th derivative $a_{PHY}(\lambda)$ marker features highlighted in Table 35.

Chlorophyte trough locations (nm)	*% of Chlorophyte spectra exhibiting this feature*	*Pigments reported as absorbing at this wavelength*
440	100%	CHLa
470	94%	CHLb
490	33%	Various xanthinols
500	72%	FCX
530	33%	FCX, PCE
545	100%	PCE
560	78%	PCE
575	100%	PCE
600	72%	PCE
620	44%	CHLa, PCE, APC
630	33%	PCC
650	78%	CHLb, APC
670	44%	CHLa
680	56%	CHLa

Table 41. The location of the 4th derivative troughs for the Cyanophyte $R_{PHY}(\lambda)$ and the percentage of the within-class spectra exhibiting that feature. Highlighted wavelengths correspond to the 4th derivative $a_{PHY}(\lambda)$ marker features identified in Table 36.

Cyanophyte trough locations (nm)	*% of spectra exhibiting this feature*	*Pigments reported as absorbing at this wavelength*
440	100%	CHL*a*
470	56%	β-CART, PRD, CHL*b*
480	44%	CART, FCX
490	44%	Various xanthinols
510	67%	PCE
520	44%	FCX
530	33%	FCX
555	78%	PCE, PCC
575	94%	PCE
585	67%	CART, CHL*c*
600	33%	PCE
610	78%	PCE
635	100%	PCE, APC
680	100%	CHL*a*

Table 42. The location of the 4th derivative troughs for the Bacillariophyte $R_{PHY}(\lambda)$ and the percentage of the within-class spectra exhibiting that feature. Highlighted wavelengths correspond to the 4th derivative $a_{PHY}(\lambda)$ marker features identified in Table 37.

Bacillario-phyte trough locations (nm)	*% of spectra exhibiting this feature*	*Pigments reported as absorbing at this wavelength*
440	100%	CHLa
465	100%	CHLc, β-CART
495	94%	PCE, PUB
510	28%	PCE
520	61%	FCX
545	100%	FCX, PRD, PCE
575	78%	PCE
590	22%	CHLc, CART
605	94%	WATER
630	100%	CHLc, PCC, CART
670	100%	CHLa

There were small differences in the location of some of the peak features identified in the 4th derivative $a_{PHY}(\lambda)$ and corresponding features in the 4th derivative $R_{PHY}(\lambda)$. This must be considered when assigning pigments to $R_{PHY}(\lambda)$ trough features. Furthermore, even though this study only considered the $R_{PHY}(\lambda)$ from two species for each class, there was considerable within-class variability in the location of the 4th derivative troughs. However, this variability was by spectra rather than by species hence it can not be explained by variation in between-species pigment composition. Similarly, the variability has no correlation with CHL*a* concentration. It is more likely to have arisen from measurement noise and from the artefacts of the derivative calculation procedure.

Pure $a_{H2O}(\lambda)$ 4th derivative {4,4,29,1} peaks occur at 605, 645, 670, 695 and 710 nm corresponding to shoulders in $a_{H2O}(\lambda)$ as discussed in § 2.1.3. These features were consistently present in the Bacillariophyte $R_{PHY}(\lambda)$ and the $R_{PHY}(\lambda)$ from the other classes at the lowest concentrations. The

Bacillariophyte spectra were notable in that they never reached high CHL*a* concentration in the tank. This suggests that derivative features attributable to $a_{H2O}(\lambda)$ will only be perceptible at when phytoplankton concentrations are low.

Using the derivative feature locations compiled from the pure culture 4th derivative $a_{PHY}(\lambda)$, summarised in Table 39, the pure culture 4th derivative $R_{PHY}(\lambda)$ were subjectively 'assigned' a class. The assignment was conducted blind without prior knowledge of the actual class of each spectra. 72% (39 out of 54 spectra) were correctly assigned in this manner. The Chlorophytes were most reliably assigned, while the Bacillariophytes were least reliably assigned (being classified as Cyanophytes). Misclassifications tended to be for spectra at the lower CHL*a* concentrations. However, the $R_{PHY}(\lambda)$ for *Asterionella* were misclassified as Cyanophytes at both low and high CHL*a* concentrations because of the occurrence of a derivative feature at 630 nm which was taken to be indicative of Cyanophyte.

7.3.5 Discussion of the derivative analysis technique

The findings of this research were similar to those of Richardson *et al.* (1994). Using 4th derivatives of reflectance spectra measured with a Spectron SE-590, they found that CHL*a* and some of the phycobilins were detectable. In particular the spectral features of PCC at 620 nm and MYX at 505 nm were identifiable, with these pigments being used to correctly diagnose the presence of a Cyanophyte bloom. Their derivative analysis was corroborated by HPLC analysis (after Mantoura and Llewllyn 1983) of actual pigment concentrations. Their approach was flawed in that they were basing their pigment identification on the location of pigment absorption spectra as measured from pigments extracted in organic solvents. Extracting pigments in solvents modifies their absorption properties leading to hypsochromatic and bathychromatic shifts in the location of the peak absorbance. Not taking these shifts into account increases the chance of a mis-classification of a pigment. Furthermore, they identified some derivative features as being atmospheric features. Had the reflectance been correctly calculated then atmospheric features should not have been present in the ground truth reflectance spectra used in their analysis.

This study has confirmed that derivative analysis can be successfully applied to the problem of identifying phytoplankton class from both $a_{PHY}(\lambda)$ and $R_{PHY}(\lambda)$. Using the 4th derivative features in Table 39 to classify the mixed culture $a_{PHY}(\lambda)$ a successful identification of 90% of either of the two classes was achieved, with a 60% success rate at the identification of the dominant class, but only 42% success rate at identifying the sub-dominant class. When the features in Table 39 were used to classify the $R_{PHY}(\lambda)$ for pure cultures 72% were successfully assigned. These results are encouraging. They suggest that the features listed in Table 39 effectively characterise the between-class differences manifest in 4th derivative $R_{PHY}(\lambda)$ which have been shown to be caused by differences in between-class pigment composition.

The pigments identified by calculating the 4th derivatives of $a_{PHY}(\lambda)$ and $R_{PHY}(\lambda)$ tend to be chlorophylls and biliproteins, there is a marked absence of carotenoids with the exception of the xanthinols. This finding was similar to that of Richardson *et al.* (1995). This absence is for two interrelated reasons. Firstly, there are many more carotenoids than there are chlorophylls and biliproteins and they all tend to have broad $a_{PIG}(\lambda)$ bands centred in the same spectral region from 450 to 550 nm (Bidigare *at al.* 1989). This makes the carotenoids hard to identify either in raw or 4th derivative spectra. Secondly, as a consequence of the difficulty in measuring the *in vivo* $a_{PIG}(\lambda)$ for the carotenoids, there are few published references describing their centres of absorption and thus they are not well represented in the pigment database. However, the relative absence of carotenoids has not to been a problem because, again with the exception of the xanthinols, carotenoids tend not to be class marker pigments.

Applying derivative analysis to pure culture $a_{PHY}(\lambda)$ led to the identification of features which could not be ascribed to pigments according to the previously published $a_{PIG}(\lambda)$ properties as collated in the pigment database. Given that this analysis relies upon the assumption that peak features in the 4th derivative of $a_{PHY}(\lambda)$ are the centres of pigment absorption bands and assuming that these features are not pseudo derivative features (§ 7.3.1) then they must be the centres of previously unreported pigments or previously unknown absorption features of known pigments. For example, there are some Chlorophyte, Bacillariophyte and Euglenophyte $a_{PHY}(\lambda)$ 4th derivative features listed in Table 34, Table 35 and Table 38 that suggest they contain biliproteins. However, there are no literature references to the presence of biliproteins in either of these classes. The consistent within-class presence of these particular derivative features suggests that they are not pseudo features but are the result of pigment absorption. There is considerable scope for features to be either previously unreported pigments, previously unreported absorption properties of known pigments, pigments associated with the class which have undergone some form of photoadaptation or they may indeed be biliprotein pigments. Given the absence of any corroborative evidence regarding pigment composition it is impossible to substantiate any of these hypotheses.

The ability to detect a pigment's presence in a 4th derivative $R_{PHY}(\lambda)$ relies on a range of conditions being met. The pigment must have an absorption spectra which peaks at a distinctive wavelength location which is separate from the absorption centres of other pigments. The degree of spectral separation of these pigment absorption peaks is dependent upon the relative absorption efficiencies and concentrations of the two pigments. A strongly absorbing pigment at low concentrations will have a less significant $a_{PIG}(\lambda)$ than that of a weakly absorbing pigment at high concentration. Information on pigment concentration and pigment absorption efficiencies was not available during this study so no assessment of how these factors affect the derivative analysis could be performed. Theoretically absorption efficiency of a pigment is also dependent upon the pigment packaging (through the package effect as described in § 2.1.5.2). It is therefore likely that the cell morphology (internal and

external) will play an important role in determining the ability of derivatives to resolve pigment absorption features. This is particularly relevant for the Cyanophytes which do not have their pigment confined in the thylakoid packages and are consequently likely to be more efficient absorbers.

Derivative analysis also relies upon published data for the identification of pigment absorption sites. Such information cannot be corroborated by this study and cannot be considered complete. Locational discrepancies between the reported absorption wavelength for a pigment and the wavelength of the derivative feature were experienced by Bidigare *at al.* (1989) who supported their derivative analysis of pigments with separate pigment qualification and quantification through HPLC. This study would have benefited from a similarly separate accurate pigment identification routine such as HPLC allowing the confirmation of both the presence and concentration of the pigments identified as features in the derivative spectra.

In the approached used here, there is subjectivity present at a number of the steps in the derivative analysis. The choice of Savitzky-Golay parameters can influence the number and location of derivative features. The visual identification of derivative features at 5 nm intervals was also subjective, especially when the feature concerned was a subtle shoulder or when the input spectrum was particularly noisy, creating many small magnitude features in close proximity. Feature identification can easily be automated but a rigorous identification of features would tend to omit the shoulder features which are obvious when visually assessing derivative spectra and would include all features regardless of whether they were real, pseudo or noise features. The final area of subjectivity is the assignment of class on the basis of feature location, this is done by reference to published data on the class marker features. However, in derivative spectra of natural phytoplankton assemblages it is unlikely that all the necessary class marker features will be identifiable especially from the sub-dominant classes. Furthermore there are frequently 'surplus' derivative features not used in the classification which serve to confuse the analysis. Given all these subjective factors it is evident that some expertise is necessary to use derivative analysis successfully for the purpose of phytoplankton class identification from absorption and reflectance.

For the application of derivative analysis for phytoplankton identification in the field the acquirement of the hyperspectral data are necessary. Traditional sensors measure in broad bands precluding the application of derivative analysis. However the newer airborne sensors *e.g.* the AVIRIS and the CASI instruments, do have hyperspectral capabilities and so can be evaluated using derivative analysis. Indeed when Richardson *et al.* (1995) applied the technique to AVIRIS data it yielded promising results.

7.3.6 Using derivatives to predict pigment concentration

This study is primarily concerned with identification of phytoplankton class from $R_{PHY}(\lambda)$ which can be achieved through the qualification of pigment composition. However this gives only half the information required by a water manager, as information on the presence of particular classes would be enhanced by quantification of the concentrations of those classes. Other workers have used derivative analysis of $a_{PHY}(\lambda)$ for both pigment qualification and quantification in conjunction with HPLC analysis for the purposes of confirming both (Butler and Hopkins 1970, Faust and Norris 1985, Bidigare *at al.* 1989 and Smith and Albert 1994). They have shown that both qualification and quantification of the pigment composition from *in vivo* $a_{PHY}(\lambda)$ is possible. Faust and Norris (1985) and Bidigare *at al.* (1989) both publish regression equations for the purposes of quantifying pigment concentration from derivative values at specific wavelengths. The former empirically sought the wavelengths with the strongest relationship with the concentration of the pigment under consideration while the latter described the relationship between the pigment concentration and the derivative value at the wavelength where that pigment is known to absorb. The two procedures suggest different observational wavelengths for quantifying the same pigment.

In accordance with the work of Bidigare *at al.* (1989) the relationship between concentration of CHL*a* (the only pigment for which concentration was separately determined during this study) and the 4th derivative at 676 nm was assessed using the pure and mixed culture absorption spectra. The CHL*a* concentration was calculated with reasonable accuracy from the 4th derivative $a_{PHY}(\lambda)$ according to:

$$\text{CHLa } (\mu g/L) = 213463.86\ (\delta^4\ a_{PHY}(676) / \delta^4\ \text{nm}) - 369.93\ (R^2 = 0.823) \quad \text{.....7.3.6.1}$$

This confirms that certain pigment concentrations can be accurately predicted using derivative analysis. The relevance of this being that an ability to accurately measure class marker pigment concentrations may well facilitate a more quantitative assessment of bloom compositions.

7.3.7 Summary analysis of the derivative classification routine

Through its ability to enhance spectral components derivative analysis has proven itself to be a successful and rapid means of qualifying the pigment composition of cultured phytoplankton $a_{PHY}(\lambda)$. The high degree of within-class uniformity in the $a_{PHY}(\lambda)$ arising from the similarity of within-class pigment composition means that the locations of 4th derivative $a_{PHY}(\lambda)$ features also have a strong within-class similarity. These features can be used to classify $a_{PHY}(\lambda)$

Using the pure culture $a_{PHY}(\lambda)$ 4th derivative, class marker features have been identified for the purpose of classifying mixed culture $a_{PHY}(\lambda)$ and, more importantly for remote sensing, pure culture $R_{PHY}(\lambda)$. Both types of spectra were correctly classified with a high level of accuracy. The tank environment in which the $R_{PHY}(\lambda)$ were measured is essentially a Case I water. Given the nature of derivatives it would seem likely that derivative analysis could be applied to both natural water bodies

of both Case I and II optical quality. An added benefit of derivative analysis is that 4th derivatives may also be used to predict pigment concentration.

While 4th derivatives do not unequivocally identify phytoplankton pigments they do provide the potential to extract increased amounts of information over-and-above that provided by *in vivo* absorption or hyperspectral reflectance spectra thereby facilitating the rapid characterisation of pigment composition leading to classification of phytoplankton classes. It should be noted that application of derivative analysis requires a precise, consistent and high-spectral resolution data set, and that there are subjective steps in the classifying procedure requiring some level of expertise. Further development of the derivative analysis technique would benefit from the more robust but time-consuming pigment identification analysis provided by HPLC (Mantoura and Llewllyn 1983).

7.4 Using neural networks to classify phytoplankton reflectance spectra

Traditional classifications, like the canonical discriminant analysis, rely upon assumptions about the underlying statistical distribution of the data. Such parametric classifications are accurate when the data set to be classified conforms with those assumptions, but are inappropriately applied when the data set differs from statistical distributions on which the original analysis were performed (Paola and Schowengerdt 1995, Foody 1996). Neural networks are non-parametric classifiers and so have the potential to be more robust especially when knowledge of the underlying data distributions is absent.

In their review of the neural network applications in remote sensing Paola and Schowengerdt (1995) found examples of classifications comparing neural networks to conventional classifiers which suggested that in many instances they improved classification accuracy. They also found reports that smaller training sets were more accurately classified by neural networks which is important if the availability of training data sets is limited. Furthermore, an important limitation of traditional classification routines is that their output is 'hard', or definitive, compared to the 'fuzzy' or conditional output from neural networks (Foody 1996). Hard classifications are potentially wasteful of information pertaining to the strength of class membership as they were developed for classifications that were considered to be discrete and mutually exclusive. It is frequently the case that freshwater reflectance data contain signals from a mixture of classes and it would be extremely useful to water managers were the relative contributions of those classes made available. Neural networks can accommodate such multi-signal data sets because assessment of multiple or partial class membership is fundamental to fuzzy classifications (Foody 1996). Neural networks are also capable of coping with noisy data. This section describes the application of neural networks to the problem of classifying freshwater phytoplankton from reflectance spectra.

7.4.1 Neural networks: a brief introduction

It is not the purpose of this section to fully review neural networks, however a brief introduction is necessary. More complete introductions to neural network applications in remote sensing are given in Eberhart and Dobbins (1990), Paola and Schowengerdt (1995) and Foody (1996).

Neural networks are composed of input and output layers between which lie one or more the hidden layers (Paola and Schowengerdt 1995). Each layer has a number of processing nodes which are fully interconnected with all the nodes in the following layer but there are no interconnections within a layer. The connections for most networks are restricted to one-way only, such networks being known as feed-forward networks. The number of nodes in the input and output layers is determined by the structure of the input data set and the number of required output classifications. The number of nodes in the hidden layer is user-defined and is usually determined by trial-and-error. The nodes in the hidden layers are centres for processing; no processing is performed in the input layer and the output nodes act only to output information to the user. The processing involves summing inputs from the various connections and passing the results through an activation function to produce the node's output. The activation function controls the threshold at which a node will pass on an output. Processed values that pass the activation thresholds of the nodes are sent via the interconnections to the nodes in the next layer. Activation functions are typically sigmoidal because the non-linearity has been found to enhance network predictivity. The interconnections are also weighted, thereby further modifying the node's output prior to it being received by the nodes in the next layer.

There are two stages involved in using neural networks for classification: training, followed by classification (Paola and Schowengerdt 1995). Training is performed on a training data set comprising input data and their known output classifications. The goal of the training is to minimise the overall error between desired and predicted output from the network by altering the weights of each node. Training is typically performed by error backpropagation algorithms which are fully described by Rummelhart *et al.* (1986). Initially weights are randomly set and repeated iterations of the backpropagation routine are used to minimise the error in network prediction using the backpropagation algorithm to alter the weights. Alterations to the weights changes the sequence in which individual nodes reach their activation thresholds thereby modifying the final prediction. To determine how the weights are changed the error in prediction is passed back from the output layer. Changes in weights are proportional to the previous changes in weights. It is this carry over of terms from one weight change to the next that forms the heart of the backpropagation algorithm.

The training of networks typically involves obtaining a training data set, defining the network structure (number of input, hidden and output nodes), and setting the learning rate and convergence factor. The learning rate sets restrictions on the magnitude of change that can be instigated by the backpropagation algorithm. The convergence factor determines the level of accuracy to which the

network is finally trained and is important because excessive training of a network can lead to the network 'memorising' the training data set thereby reducing its ability to predict data sets other than the training data set (Paola and Schowengerdt 1995).

The drawback of feed-forward backpropagation neural networks is that their classifications are essentially performed in a 'black-box'; consequently there is a loss in understanding of how the classification was actually established due to departure from the assumption that classes have statistically well-defined distributions (Paola and Schowengerdt 1995).

Neural networks have already been employed in a number of areas dealing with the classification and general data processing of remote sensing data (Smith *et al.* 1995, Lewis *et al.* 1995, Foody *et al.* 1995, Stephanidis *et al.* 1995). Neural networks have also been used in limnological sciences, for example Boddy *et al.* (1994) used neural networks to identify marine phytoplankton species from measurements of flow-cytometric data.

7.4.2 Neural network methods

The neural network used in this study was a package called *NeuroShell* v4.1 (Ward Systems Group, Inc.). It is a feed-forward backpropagation neural network which giving a degree of flexibility in setting-up the network architecture and controlling the learning rates. *NeuroShell* can be operated in either binary or analog mode depending upon the nature of the input data.

At the outset a number of potential problems to the successful application of neural networks to the classification of $R_{PHY}(\lambda)$ were identified:

1. $R_{PHY}(\lambda)$ changes both with phytoplankton concentration and with phytoplankton class (§ 4.3.4) and those spectral changes ensuing from variation in biomass could prevent the successful application of a neural network. No suitable $R_{PHY}(\lambda)$ normalisation technique has been developed to successfully remove the biomass influence on $R_{PHY}(\lambda)$ in the manner that calculating the specific absorption spectra has done for removing biomass influences from $a_{PHY}(\lambda)$;
2. The $R_{PHY}(\lambda)$ data set is restricted in species variation with only two species from each class (§ 4.3.4). This restriction could potentially reduce the predictive capacity of the trained network;
3. Inputting training data at the high spectral resolution of the $R_{PHY}(\lambda)$ measured in this study would require an extremely untraditional network structure with up to 200 input nodes and three output nodes. The affect of such a network architecture on the network's prediction capabilities would require investigation.

One solution to these problems is essentially an extension of the derivative analysis (§ 7.3). $a_{PHY}(\lambda)$ peaks and $R_{PHY}(\lambda)$ troughs are located on the centres of absorption for pigments and these absorption sites are not susceptible to shifts in location with biomass concentration as are reflectance peaks. The

ability to use the more extensive $a_{PHY}(\lambda)$ data set in combination with the restrictive $R_{PHY}(\lambda)$ data set meant that the neural network training would be more robust.

The locations of 4th derivative {4,4,29,1} peaks in the 88 pure culture $a_{PHY}(\lambda)$ data set were identified at a resolution of 5 nm according to the method in § 7.3.3.1. These peak locations were then converted into binary form for input into *NeuroShell*. Thus the input data set reduced the hyperspectral $a_{PHY}(\lambda)$ to a series of 4th derivative peak locations by giving the value zero to wavelengths with peak absence and one for peak presence. This conversion was performed for 53 wavelengths at 5 nm intervals from 440 and 700 nm (the locations of the Soret and far-red CHL*a* absorption bands between which the other photosynthetic pigments absorb). Each data input line was appended with a class identifier which consisted of five values of four zeros and a single one, with the placement of the one indexing the class of that particular training spectra. An example of this input data is given in Figure 77.

Spectra 1	10000100001000000000100000001000000010000010000010000	10000
Spectra 2	10000000010000000000100000001000000010000100000010000	10000
Spectra 3	10000010000010000000001000100100000001000100000010000	10000
Spectra 4	10000010000010001000100100100010001000100000000100000	00010
Spectra 5	10000010000100000000100000001000000010000000000010000	00001
Spectra 6	10001001000010100010001000100010000000100000001000000	00100
Spectra 7	01000010000010000000001000000010000000000100000010000	00100
Spectra 8	10000010010000000100001000101001001000010000000010000	01000
Spectra 9	10000010101000001000001000011010001000001000000010000	01000
.....	....	.....
.....	....	.....
Spectra 88	10010010100000001000001000011010001000001000000010000	01000

Figure 77. Example of the 4th derivative peak location $a_{PHY}(\lambda)$ input training data used in this study. A value of 1 at position *n* in the first 53 digits of each spectra symbolises a peak at wavelength (400 + 5*n*) nm. The last five digits are used to convey phytoplankton class: 10000 = Chlorophyte; 01000 = Cyanophyte; 00100 = Bacillariophyte; 00010 = Cryptophyte; 00001 = Euglenophyte.

The network architecture comprised three layers with 53 input nodes, the default setting of hidden nodes and five output nodes. For *NeuroShell* the default setting of hidden nodes is two times the square root of the number of input plus output nodes, so in this instance is 15. The other user specified options during *NeuroShell* training were:

1. Learning threshold - the convergence factor, which determines the accuracy of the trained network. When the error level of all of the input data are less than this threshold the training automatically ceases;

2. Learning rate - determines the maximum size of changes in weights that can be performed during the backpropagation training routine;
3. Momentum - determines the slope of the activation curve. Using a high momentum in the initial stages of training prevents the network becoming 'stuck' in local minima. As training progresses a lower momentum can be used to 'fine-tune' the network.

The learning threshold was initially set to 0.01. The learning rate and momentum factors were set high at the start of training and lowered as the training progressed.

7.4.3 Neural network training

Training was completed in 10,000 iterations. At this relatively small level of training the network was able to correctly identify the class of 99% of the input spectra $a_{PHY}(\lambda)$ (87 out of 88). Figure 78 shows the error factors for each of the 88 spectra computed by the trained network. Error factors are simply the difference between what the network predicts for a spectrum and the actual value. For most of the training spectra the error factors were low.

Input spectra	Error factors				
1 through 5	0.00785	0.00768	0.00498	0.00835	0.00516
6 through 10	0.00669	0.00963	0.00756	0.00652	0.00950
11 through 15	0.00879	0.00866	0.00194	0.00838	0.00242
16 through 20	0.03905	0.00920	0.00268	0.00682	0.00770
21 through 25	0.00968	0.00127	0.00284	0.00270	0.00990
26 through 30	0.00795	0.00713	0.00631	0.00592	0.00956
31 through 35	0.06126	0.00469	0.00304	0.00822	0.00773
36 through 40	0.00778	0.00465	0.00406	0.00802	0.00847
41 through 45	0.00934	0.00772	0.01434	0.00655	0.01809
46 through 50	**0.33313**	0.00993	0.00823	0.00688	0.00995
51 through 55	0.00986	0.00806	0.00903	0.00694	0.00833
56 through 60	0.00964	0.00443	0.00294	0.00878	0.00344
61 through 65	0.00390	0.00872	0.00540	0.00563	0.00723
66 through 70	0.00410	0.00419	0.00470	0.00627	0.00424
71 through 75	0.00456	0.00571	0.00675	0.00623	0.00560
76 through 80	0.00407	0.00630	0.00695	0.00423	0.00413
81 through 85	0.00688	0.00462	0.00445	0.00550	0.00982
86 through 88	0.00389	0.00428			

Figure 78. The learning error factors by input spectra for the network trained with the full set of 4th derivative pure culture $a_{PHY}(\lambda)$ peak locations. The learning error factors shows the actual error factor for each sample case which is based on the differences between what the network predicted and the actual values. The smaller the error the more accurate the prediction. The high error value of the single mis-classified Euglena $a_{PHY}(\lambda)$ is highlighted in bold.

Predictions given were in the form of a five scores, one for each class used in the training data set, which ranged from zero to one indicating the likelihood of the spectra belonging to either of the five input classes. For the majority of spectra this identification was absolute, or positive, with the network assigning values > 0.8 to the correct class of the spectra and values < 0.2 to the other classes. The only incorrectly assigned spectrum was one of the *Euglena* spectra which

scored 0.66 for the class Chlorophyceae, < 0.12 for the classes Cyanophyceae, Bacillariophyceae and Cryptophyceae, and 0.35 for the class Euglenophyceae. The classification for this spectrum was both incorrect and weak because the difference between the highest and second highest scores was small, indicating that no single class was definitively identified. Two correctly classified *Chlorella* spectra were weakly assigned with the Euglenophyte score being almost as strong as the Chlorophyte.

Allowing the training to proceed to 20,000 iterations (by reducing the Learning threshold) only slightly improved the prediction accuracy but still failed to correctly identify the previously misclassified *Euglena* spectrum. Because this particular *Euglena* spectrum was consistently misclassified it was removed from the training data set and the network re-trained. This was an acceptable practice because a network trained with 'incorrect' data would have reduced accuracy when used for prediction. Further investigation of this particular *Euglena* absorption spectrum showed it to have a 4th derivative peak at 640 nm which was not present in the other Euglenophyte spectra and this appeared to be the main source of the network error.

The training of the network without the single *Euglena* spectra was performed to a Learning threshold of 0.1 (set higher to prevent over learning by the network). Training was completed in 2,350 iterations. This trained network was then used to classify the mixed culture $a_{PHY}(\lambda)$ and the pure culture $R_{PHY}(\lambda)$ data sets.

7.4.4 Results of the neural network classification

7.4.4.1 Classification of the pure culture absorption spectra

The trained network proved 100% successful at classifying the pure culture $a_{PHY}(\lambda)$. The success of the network is attributable to the fact that peaks in $a_{PHY}(\lambda)$ represent the centres of absorption for pigments, therefore the peak location data (as shown in Figure 77) acted as a 'bar-code' for each spectrum. The network was able to identify the within-class commonalties and between-class differences and use this information as a basis for predictions. The classification of the Chlorophytes and Euglenophytes was not as positive as the classification of the other classes. These two classes have already been highlighted as having very similar pigment compositions (§ 4.1) and their apparent spectral similarity would be greater at this relatively low spectral resolution. However, it could be argued that these exceptions 'prove the rule' and pigment differences between these particular classes may require higher resolution to be positively identified.

It has already been noted that the use of a backpropagation training algorithm leaves the user with little on which to assess the internal structure of the network. However *NeuroShell* gives the user the ability to view the weights applied to the nodes in the trained network (each node representing a wavelength). This is useful because it allows the user to check the contribution each node has upon the

network's output which is summarised in the 'Contribution Factor'. The Contribution Factor is a relative measure of the connection weights from each input to every hidden node and is shown for each wavelength in the trained network in Table 43. Those wavelengths that contribute most to the successful training of the network have the highest Contribution Factors.

It is not surprising to find that wavelengths with the highest Contribution Factors are those that have been identified throughout this study as being the centres of absorption for the class marker pigments. Notably 650 nm (CHL*b*), 620 nm (PCC), 630 nm (PCE) and 520 nm (FCX). Table 43 shows that 495 and 475 nm have high Contribution Factors, these wavelengths have not previously been noted as being important marker pigment absorption wavelengths. Pigments reported as absorbing at 495 nm are PCE and PUB, and at 475 nm are CHL*b* and FCX. The network's pattern recognition of the 4th derivative 'bar-codes' is undoubtedly more complex than simple identification of the presence of single marker pigments as was performed during the derivative analysis study. The network is probably making a more comprehensive use of all wavelengths as indicated by the strength of the Contribution Factors for many of the other nodes. For instance, 680 nm has a high Contribution Factor which could be the network using the location of the far-red CHL*a* peak to distinguish between-class with the Chlorophytes, Cyanophytes and Euglenophytes having far-red CHL*a* peaks at longer wavelengths and Bacillariophytes and Cryptophytes at shorter wavelengths. It will also be noted that there are wavelengths whose Contribution Factors are very small indicating a dearth of information relative to the identification of phytoplankton class. Wavelengths longer than the far-red CHL*a* absorption band and wavelengths in close proximity to the Soret absorption band all have low Contribution Factors suggesting that they contain little relevant information to the problem of class identification.

Table 43. Contributions Factors (ordered by rank) for each input wavelengths (nodes) in the neural network trained on the locations of peaks in the 4th derivative pure culture $a_{PHY}(\lambda)$. The Contribution Factor describes the relative importance of a particular node to the predictions of the trained network.

Wavelength (nm)	650	620	680	495	475	630	520	635	645
Contribution	31.7	25.4	23.2	22.4	22	20.6	20.4	20.2	19.8
Wavelength (nm)	485	550	470	610	445	515	540	530	675
Contribution	19.3	19.2	18.9	18.7	18.2	17.6	17.2	17.1	16.4
Wavelength (nm)	640	545	585	570	460	500	670	590	565
Contribution	16.3	15.6	15.5	15.3	14.7	14.4	13.7	13.6	13.3
Wavelength (nm)	440	660	465	625	555	580	510	525	490
Contribution	13	13	12.7	12.7	12.4	12.4	12.2	12.2	11.6
Wavelength (nm)	560	600	480	685	615	575	655	505	605
Contribution	11.1	11.1	10.2	10.2	9.8	8.5	7.2	6.4	6.0
Wavelength (nm)	455	535	695	595	665	690	700	450	
Contribution	5.4	5.0	2.5	2.4	2.4	1.9	1.9	1.6	

7.4.4.2 Classification of the mixed culture absorption spectra

The network trained on the pure culture $a_{PHY}(\lambda)$ was then used to identify the dominant class in the fifty mixed culture $a_{PHY}(\lambda)$. The network correctly identified the dominant class in 70% (35 out of 50) of spectra. If success was measured as identification of either one of the two species in each mixture then the network correctly identified 82% (41 out of 50) of the spectra.

The network correctly identified the dominant class in 90% (25 out of 28) of those spectra where one class comprised > 75% of the total CHL*a* concentration in the mixture. Of the remaining spectra, where neither class was excessively dominant, only 45% (10 out of 22) of the dominant classes were correctly identified. The network is therefore better at identifying the dominant class of a mixed culture $a_{PHY}(\lambda)$ when one class is very much more concentrated than the other.

The network classifies spectra by assigning a score, the class with the highest score being deemed the dominant class. The mixed $a_{PHY}(\lambda)$ assignments tended to be strong, this was the case regardless of whether the classification was correct or incorrect, and regardless of the relative concentrations of the actual classes in the mixture. Out the outset it was hoped that the network would indicate the relative proportions of the dominant and sub-dominant classes however there was no evidence of the reported fuzzy prediction capabilities of neural networks (Foody 1996). The absence of any successful fuzzy predictions is possibly because training a neural network in this manner required the network to use binary input data as opposed to analog data. Had it been possible to use analog input data the fuzzy predictions may have become more apparent.

7.4.4.3 Classifying the pure culture reflectance spectra

The network trained using the peak locations of the 4th derivative pure culture $a_{PHY}(\lambda)$ was then used to classify the pure culture $R_{PHY}(\lambda)$ spectra which had been likewise converted, this time using 4th derivative trough locations. The network correctly identified 81% of the $R_{PHY}(\lambda)$, the classifications for each individual spectrum in the pure culture $R_{PHY}(\lambda)$ series are given in Table 44.

Table 44. Classification of the 54 pure culture phytoplankton reflectance spectra using the neural network trained on the pure culture phytoplankton absorption spectra. The network classifications for the nine spectra for each phytoplankton species are denoted by an index: 1= Chlorophytes; 2 = Cyanophytes; 3 = Bacillariophytes; 4 = Cryptophytes. The index for the actual class of each species is shown in brackets. Spectra that were misclassified are highlighted in bold.

Phytoplankton species	*Classification by series number for each pure culture reflectance spectra*								
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
Selenastrum (1)	**4**	**4**	**4**	**4**	1	1	1	1	1
Chlorella (1)	**4**	1	1	1	1	**3**	**3**	1	1
Anabaena (2)	2	2	2	2	2	2	2	2	2
Microcystis (2)	**4**	2	2	2	2	2	2	2	2
Asterionella (3)	3	3	3	3	3	3	**4**	3	3
Fragilaria (3)	3	**4**	3	**4**	3	3	3	3	3

Of the three classes the Cyanophytes (*Anabaena* and *Microcystis*) and Bacillariophytes (*Asterionella* and *Fragilaria*) were most consistently identified correctly. Identification of the Chlorophytes (*Selenastrum* and *Chlorella*) was less successful. Three *Selenastrum* $R_{PHY}(\lambda)$ (the second, third and forth concentrations in the series) were positively misclassified as Cryptophytes. Three of the *Chlorella* $R_{PHY}(\lambda)$ (the first, sixth and seventh) were misclassified as either a Bacillariophytes or Cryptophytes. However, for the seventh scan in the concentration series, this misclassification was weak with network scores for Chlorophytes and Cryptophytes being 0.45 and 0.48 respectively. All nine of the *Anabaena* $R_{PHY}(\lambda)$ were correctly and positively identified. The first spectrum of the *Microcystis* $R_{PHY}(\lambda)$ series was misclassified as a Cryptophyte and the seventh spectrum of the *Asterionella* $R_{PHY}(\lambda)$ series was misclassified as a Cryptophyte. Two *Fragilaria* $R_{PHY}(\lambda)$ (the second and forth) were misclassified as Cryptophytes, however, the latter was a weak misclassification with Bacillariophytes scoring 0.71 and Cryptophytes scoring 0.77.

7.4.5 Discussion of neural network classification

These results show that a network trained using the 4th derivative peak locations of the pure culture $a_{PHY}(\lambda)$ can be used to classify pure culture $R_{PHY}(\lambda)$ 4th derivative trough locations. The use of neural networks should therefore be applicable to the problem of classifying reflectance spectra from natural Case I freshwater bodies. The presence of strong concentrations of tripton in Case II waters may reduce the success of such a classification. The presence of coloured DYS should not have an effect upon network prediction success because the exponential nature of $a_{DYS}(\lambda)$ will be removed on calculation of the 4th derivative.

The application of a neural network to the problem of identification of phytoplankton from reflectance spectra has proved successful. The technique employed here is essentially a more objective extension of the derivative analysis described in the previous section. Training a network on a data set comprised of $a_{PHY}(\lambda)$ 4th derivative peak locations (which correspond to the centres of pigment absorption) has allowed the network to develop a pattern recognition routine for the classes. Having developed this ability to interpret the 4th derivative $a_{PHY}(\lambda)$ 'bar-code' the process can be reversed and the classes from 'unknown' $a_{PHY}(\lambda)$ or $R_{PHY}(\lambda)$ can be identified.

The precision of the network's ability the classify both the reflectance and the absorption spectra may have been improved had the spectral resolution been improved. The spectral resolution for the input spectra was reduced from 2 nm to 5 nm, furthermore the assessment of a 4th derivative peak and trough features at each wavelength was, to a small degree, subjective. The technique may well have been improved had the peak identification been automated. Increasing the spectral resolution to 2 nm would have increased the amount of spectral information presented to the network for training, however, it would have created a very unconventional network structure. The training of such a network may well have been difficult. Moreover, the Savitsky and Golay derivative calculations use a

smoothing factor which will remove some of the spectral detail from the final derivative spectrum. Further work would be necessary to assess whether a 2 nm resolution would successfully improve upon the classification potential of the neural network.

A drawback of the application of neural networks to the problem of spectral classification is that it is not possible to extract the basis for the pattern recognition from the network. There is no comprehensive analytically based understanding of why a network is performing effectively nor of those instances when it may fail. This application has assumed that because the network has been trained upon absorption spectra from phytoplankton classes which contain different accessory pigments that the unique absorption features associated with these pigments has been the basis for the network's classification scheme. This study has incorporated an assessment of the Contribution Factors (Table 43) which have indicated this assumption to be correct.

It should also be noted that using input spectra converted to a derivative for the training and classification of this neural network subjects the final classification to the problems associated with the calculation of the derivative spectra. These have been dealt with in the previous section.

7.5 Conclusions for the phytoplankton class identification routines

Four routines for the identification of phytoplankton class from reflectance spectra have been presented. Although they employ different techniques they are all based upon the same approach: the identification of class marker pigments. This has been achieved through the identification of the unique absorption features exhibited by these marker pigments which are manifest in the phytoplankton reflectance spectra. Airborne remote sensing systems can therefore be used to identify the dominant phytoplankton type in a target water.

The four routines have different requirements for the spectral resolution of the input data. The first two routines only require input reflectance data in approximately 10 nm wide bands, while the derivative and neural network routines require hyperspectral reflectance spectra at 2 nm. Consequently the latter two are only applicable to the most modern of airborne remote sensing systems such as the CASI sensor (flown in hyperspectral mode).

Identification of reflectance ratios to distinguish between phytoplankton classes has never been attempted before. The F-ratio matrix was shown to be an effective means for identifying the absorption and reflectance ratios with the strongest propensity to discriminate between two phytoplankton classes. The class identification ratios (Table 21) developed, from the pure culture $a_{PHY}(\lambda)$ spectra, were successful at discriminating between two classes of phytoplankton. This discrimination was highlighted by box plots (Figure 58 to Figure 60). These ratios have drawn upon observational wavelengths where class distinctive marker are known to absorb. The rationalisation of

these absorption ratios (Table 22) has highlighted five significant class marker pigments at five observational wavelengths. These pigments were identified in the literature as being the most likely class distinctive marker pigments (Table 1).

The class ratio technique was difficult to transpose to enable it to be applied to reflectance spectra. The lack of correspondence between the absorption and the reflectance class identification ratio wavelengths, and the lack of correspondence between reflectance ratio observation wavelengths and the absorption centres for class marker pigments, meant that the technique's success was difficult to explain. The restrictions of discriminating between only two pre-selected classes also reduces the operability of this technique. The most useful outcome of the class ratio technique has been the confirmation of the thesis that the optical properties of class marker pigments can be used to discriminate between phytoplankton classes. As such the class ratio identification technique is less a predictive tool and more an analytical tool for the understanding of phytoplankton optics and identifying wavelengths which are important in distinguishing between phytoplankton classes.

The dimension reduction capabilities of the canonical discriminate analysis technique is its most obvious characteristic. The technique identified a handful of wavelengths from hyperspectral data with the strongest propensity to discriminate between combinations of up to five classes of phytoplankton. The thesis that class marker pigments would facilitate the identification of classes in high resolution spectra was confirmed. The technique selected wavelengths associated with class marker pigments (Table 26). Those class marker pigments responsible for the selection of these observational wavelengths were those pigments identified in Table 1 as being pigments most likely to be class marker pigments. The plots of the canonical scores for phytoplankton absorption (Figure 66) and reflectance spectra (Figure 69) in canonical space highlight the ability of the canonical discriminant analysis to succinctly discriminate between phytoplankton classes.

There was little correspondence between those wavelengths selected by the canonical discriminant analysis when calculated for the phytoplankton absorption spectra and those selected for the phytoplankton reflectance spectra. Although this casts doubt upon the thesis that phytoplankton reflectance is the inverse of phytoplankton absorption this situation is likely to have arisen because of the subtly of pigment features in reflectance spectra. This highlights the difficulties of identifying phytoplankton classes from remotely sensed data. The reliance of this parametric technique upon a particular data set to calculate the canonical correlations used to classify 'new' spectra by this parametric technique will reduce the applicability of this technique to spectra not acquired under the same optical conditions as the original data set.

The benefits of using derivative analysis in remote sensing has already been proven (Demetriades-Shah *et al.* 1990). This is also true of the application of derivatives to the problem of phytoplankton

class identification (Bidigare *at al.* 1989, Richardson *et al.* 1994). In this study the application of this non-parametric technique differs slightly in its approach as the problem of noise generated by the derivative calculation has been addressed by the identification at 5 nm intervals of those 4th derivative peak features which have been shown to confirm the presence of class marker pigment in $a_{PHY}(\lambda)$.

The pure culture $a_{PHY}(\lambda)$ were used to identify where class marker features would be enhanced by calculation of the 4th derivative (Table 39). The presence of these features could then be used to identify phytoplankton class from 4th derivative $R_{PHY}(\lambda)$. Again the pigments associated with the class distinctive marker features in the 4th derivative were those pigments identified in the literature as the pigments most likely to distinguish between phytoplankton classes (Table 1).

The derivative technique proved successful in the classification of both the pure and mixed culture $a_{PHY}(\lambda)$ and of the pure culture $R_{PHY}(\lambda)$. Consequently it would appear to be a very robust technique with strong potential for being an operational technique for water managers with access to high spectral resolution remote sensing equipment. The technique can not be applied to broad band remote sensing data.

Neural networks have never been applied to the problem of using remote sensing to identify phytoplankton class. The neural network technique built upon the derivative technique and attempted to remove the subjectivity of the aforementioned technique. An advantage of the neural network technique over the derivative technique is that it provides an indication of the accuracy of the final classification. Although a robust technique, it was less successful than the derivative analysis technique in the classification of the phytoplankton spectra. To its detriment classification by neural networks is performed within a black box giving little indication of how the classification has been achieved.

All the class identification techniques presented have proved successful for classification of the pure culture $R_{PHY}(\lambda)$ measured under the controlled conditions of the experimental tank. The following chapter assesses their ability to classify the reflectance spectra from Esthwaite Water.

Chapter Eight Testing the class identification routines

The phytoplankton class identification routines were all developed on absorption and reflectance spectra from pure cultures of phytoplankton measured in the environment of the experimental tank. Under these controlled conditions they have all proved very successful at identifying the dominant phytoplankton class. The Esthwaite Water reflectance spectra provide an opportunity to test how these routines can be adapted to identify the dominant class from spectra measured from a natural water body. The results of this comparison exercise are presented below.

8.1 Identifying phytoplankton class from general spectral features

In § 4.3.4 the general between-class differences in the spectral form of the pure culture reflectance spectra were discussed. It was noted that the presence of three reflectance peaks within the visible spectrum was indicative of Cyanophytes. It was also noted that the wavelength location of the green reflectance peak had potential to be used as a simple class identification routine. Both these spectral features have the potential to be converted into simple class identification algorithms.

8.1.1 Identifying class using the number of distinctive reflectance peaks

The Esthwaite Water $R(\lambda)(-0.01\ m)$ spectra were splined to 10 nm to remove noise induced spectral fluctuations and the number of distinct reflectance peaks counted. The presence of three peaks in the spectra were taken to indicate the presence of a dominant Cyanophyte species. Spectra with less than three peaks were considered to be dominated by an unspecified non-Cyanophyte class. The predicted class was checked against the dominant class as identified in the surface water sample cell counts.

This routine successfully identified 82% (9 out of 11) of the Cyanophyte dominated spectra and avoided assigning a Cyanophyte class to any of the spectra which the cell counts indicated to be dominated by a non-Cyanophyte class. The results are summarised in Table 45.

Table 45. The results of using the number of peaks in the Esthwaite Water R(λ)(-0.01 m) to identify dominant class. Correct classification of Cyanophyte dominated spectra has been highlighted in the actual class column.

Date	*Number of peaks*	*Predicted class*	*Actual class*
06-Jul-94	3	Cyanophyte	Cyanophyte
14-Jul-94	3	Cyanophyte	Cyanophyte
28-Jul-94	3	Cyanophyte	Cyanophyte
01-Sep-94	3	Cyanophyte	Cyanophyte
21-Sep-94	3	Cyanophyte	Cyanophyte
23-May-95	2	unspecific non-Cyanophyte	Chlorophyte
06-Jun-95	2	unspecific non-Cyanophyte	Cryptophyte
20-Jun-95	2	unspecific non-Cyanophyte	Cryptophyte
04-Jul-95	2	unspecific non-Cyanophyte	Cyanophyte
25-Jul-95	3	Cyanophyte	Cyanophyte
01-Aug-95	2	unspecific non-Cyanophyte	Cyanophyte
15-Aug-95	3	Cyanophyte	Cyanophyte
12-Sep-95	3	Cyanophyte	Cyanophyte
10-Oct-95	3	Cyanophyte	Cyanophyte
10-Apr-96	2	unspecific non-Cyanophyte	Bacillariophyte

Only two Cyanophyte dominated spectra were incorrectly classified (6/7/95 and 1/8/95). The phytoplankton concentrations on the 6/7/95 and 1/8/95 were lowest and second lowest respectively for any of the Cyanophyte dominated spectra (Table 11). The cell counts for 6/7/95 revealed that although *Anabaena* was dominant it was only weakly so with *Rhodomonas* (a Cryptophyte) being sub-dominant. For both sample dates the species diversity index was relatively high. The experimental tank work has shown that when phytoplankton biomass concentrations are low the reflectance signal is weak and less spectrally featured. George and Charlton (1996) also found that oligotrophic lakes tended to exhibit reflectance spectra with less pronounced spectral features. Furthermore the reflectance signal, being the amalgamation of the weakly dominant Cyanophytes and the sub-dominant classes, is unlikely to contain sufficient information to adequately identify the dominant phytoplankton class using such a simple method. This-not-withstanding, the peak number identification technique has proven itself successful at distinguishing between Cyanophyte and non-Cyanophyte dominated reflectance spectra.

8.1.2 Identifying phytoplankton class from the location of the green reflectance peak

The location of the green (*ca.* 550 nm) peak in the Esthwaite Water reflectance spectra was then used to identify the dominant class according to the locations of the green peak identified for each class from the pure culture reflectance spectra as measured in the experimental tank (Figure 32 and Table 9). As described in § 4.3.5 the Chlorophyte $R_{PHY}(\lambda)$ had their green peak at 534 nm, Cyanophytes at 545 nm and Bacillariophytes at 567 nm. Class was assigned to the Esthwaite Water reflectance spectra on the basis of proximity of the green peak to either of these wavelengths.

Using this method all the Esthwaite Water R(λ)(-0.01 m) were incorrectly assigned the class Bacillariophyceae with the exception of 23/05/95 which was correctly assigned the class

Chlorophyceae. In an attempt to improve this classification routine the peak wavelength location for each class was 're-calibrated' according to the peak location of the Esthwaite Water spectra for 28/07/94 (which was overwhelmingly dominated by the Cyanophyte *Anabaena*) and the peak location of the class Cryptophyceae included (as assessed from the Cryptophyte dominated spectra on 6/6/95). The new green peak locations by class become Cyanophyte 565, Chlorophyte 540, Bacillariophyte 570 and Cryptophyte 580 nm. This significantly improved the success of this simple routine to 87% (13 out of 15) as summarised in Table 46.

Table 46. The results of using the green reflectance peak location to classify the dominant class from the Esthwaite Water R(λ)(-0.01 m). Spectra for which the identified dominant class matches the dominant class identified in the surface water cell counts have been highlighted.

Date	Actual dominant class	Dominant class according to the green peak location given in Table 9	Dominant class according to re-calibrated green peak location
06-Jul-94	Cyanophyte	Bacillariophyte	Cyanophyte
14-Jul-94	Cyanophyte	Bacillariophyte	Cyanophyte
28-Jul-94	Cyanophyte	Bacillariophyte	Cyanophyte
01-Sep-94	Cyanophyte	Bacillariophyte	Cyanophyte
21-Sep-94	Cyanophyte	Bacillariophyte	Bacillariophyte
23-May-95	Chlorophyte	Chlorophyte	Chlorophyte
06-Jun-95	Cryptophyte	Bacillariophyte	Cryptophyte
20-Jun-95	Cryptophyte	Bacillariophyte	Cyanophyte
04-Jul-95	Cyanophyte	Bacillariophyte	Cyanophyte
25-Jul-95	Cyanophyte	Bacillariophyte	Cyanophyte
01-Aug-95	Cyanophyte	Bacillariophyte	Cyanophyte
15-Aug-95	Cyanophyte	Bacillariophyte	Cyanophyte
12-Sep-95	Cyanophyte	Bacillariophyte	Cyanophyte
10-Oct-95	Cyanophyte	Bacillariophyte	Cyanophyte
10-Apr-96	Bacillariophyte	Bacillariophyte	Bacillariophyte

Thus, the simple peak location technique proved successful following the 're-calibration' of the class green peak locations. However these re-calibrated peak locations are notably different from the locations of the green peaks in the pure culture $R_{PHY}(\lambda)$. For the Chlorophytes and Bacillariophytes a minor shift towards the longer wavelengths was necessary but for the Cyanophytes a 20 nm shift towards the longer wavelengths was necessary before the technique became successful. These differences probably arise because the $R_{PHY}(\lambda)$ were measured in the absence of any other optically active parameter, while the Esthwaite Water spectra included all the other optically active components present in a natural water body. In particular the presence of DYS may be responsible for shifting the peak locations towards the longer wavelengths.

When applied to the CASI spectra both these simple spectral feature techniques indicated the dominance of Cyanophytes. All the CASI spectra had three peaks in the visible spectrum, and all had green peaks at 565 nm (Figure 41). The dominance of the Cyanophyceae was confirmed by the cell counts.

8.2 Application of the class identification ratios to the Esthwaite Water reflectance spectra

The class identification ratios in Table 20 were then used to classify the Esthwaite Water reflectance spectra. To test the classification, the number of cells per class were converted into a percentage of the total number of cells in the sample for the surface sample cell counts. These percentages were then used to calculate ratios describing the relative concentration of each class to each of the other classes present in the sample. Henceforth these ratios will be referred to as the cells-per-class ratios. Ratios of the natural log of cells-per-class were also considered in an attempt to normalise the distribution of cell numbers for each class. A sub-sample including only those spectra from 1995 was considered to compare the predicted class with the dominant class identified in the 5 m integrated sample's cells-per-class ratios.

Changes over time in the value of the reflectance ratios were compared with changes in the cells-per-class ratios to assess the class identification ratio's ability to predict changes in the relative contribution of the classes. Reflectance ratios which included the Euglenophytes were not considered because of the low presence of this class in Esthwaite Water.

Table 47 shows the correlation coefficients for the reflectance ratios against various sets of cells-per-class ratios. Outlying data-points caused by the 'absence' or very low presence of a particular phytoplankton class on a particular sample date were removed from the data set used to calculate these ratios.

When the full data set is considered the class identification ratios are poor predictors of change in the relative concentration of the different classes (Table 47). However the class identification ratios are better at explaining the variance in the cells-per-class ratios when the data sets are restricted to the 1995 sampling period (it will be recalled that during 1995 the Esthwaite Water reflectance measurements were performed in the experimental tank). The correlation coefficients for the class identification ratios against the cells-per-class ratios for the 1995 5 m integrated samples are stronger than those for the surface water samples. Using the natural logarithm of cell numbers in the cells-per-class ratio typically improves the ability of the class identification ratios to explain variance in the proportions of each class. The strongest relationships between the class identification ratio and cells-per-class ratio are achieved by using the integrated samples for 1995 and taking the natural log of the cell numbers prior to calculating the cells-per-class ratio. The strongest correlation coefficient was for the Cyanophyte to Cryptophyte class identification ratio which explained more than 77% of the variance in the cells-per-class ratio for these two classes (Figure 79). Increases in the value of the cells-per-class ratio and decreases in the value of the reflectance ratio indicate the dominance of Cyanophytes over Cryptophytes. Figure 79 shows that up until September 1995 the class identification

ratio was very successful at predicting the changes in the relative concentrations of the Cyanophyte and Cryptophyte cell numbers. It has therefore been shown that the class identification ratios can be used to monitor relative changes in the phytoplankton crop.

Table 47. The correlation coefficients for the reflectance ratios calculated from the Esthwaite Water R(λ)(-0.01 m) and the various cells-per-class ratios.

Two-class reflectance ratios	*Full data set surface samples*	*Full data set surface samples logged*	*1995 surface samples*	*1995 surface samples logged*	*1995 5 m integrated samples*	*1995 5 m integrated samples logged*
Chlorophyte : Cyanophyte	0.272	-0.521	0.452	-0.515	-0.600	0.661
Chlorophyte : Bacillariophyte	0.329	-0.128	0.560	0.377	0.503	-0.772
Cyanophyte : Bacillariophyte	0.163	-0.352	0.511	-0.453	-0.871	0.629
Cyanophyte : Cryptophyte	-0.347	0.513	-0.562	0.571	0.703	-0.883
Cryptophyte : Chlorophyte	-0.162	0.409	-0.015	-0.392	0.348	-0.483
Cryptophyte : Bacillariophyte	-0.384	0.625	-0.373	0.677	-0.284	0.313

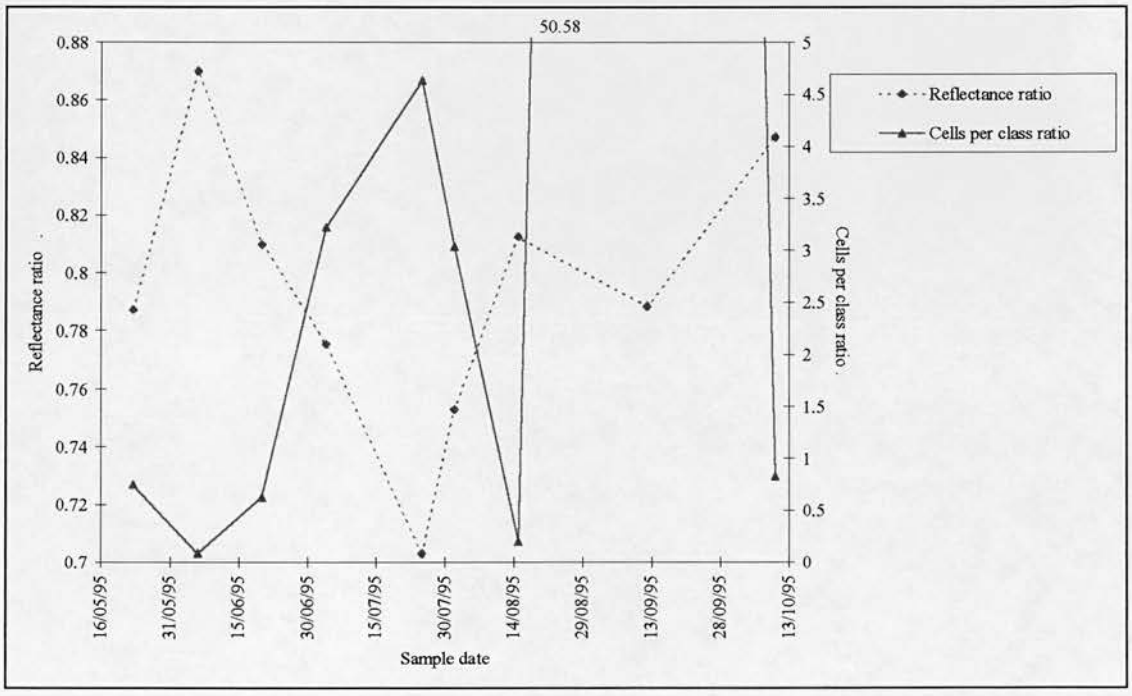

Figure 79. The Cyanophyte to Cryptophyte reflectance ratio and the log-normalised cells-per-class ratio for the integrated sample during the 1995 sampling season. Low values in the reflectance ratio and high values in the cells-per-class ratio both indicate the dominance of Cyanophytes over the Cryptophytes. Note the curves have been plotted on different axes and the data point for the cells-per-class ratio for 12/9/95 has not been plotted for display purposes.

The relationship between the reflectance ratio and the cells-per-class ratio calculated for the 5 m integrated samples is strong throughout 1995 with the exception of 12/9/95. On this sample date the dominance by Cyanophytes was overwhelming with 84% of the counted cells being Cyanophytes and only 5% being from the class Cryptophyceae.

To predict the relative concentrations of the two classes the class identification ratio algorithms were developed specifically for Esthwaite Water using the spectral signatures acquired during 1995 (Table 48). The result of each class identification ratio algorithm is an index of the dominance of one class over the other in a natural water body. This index is based upon the normalised cells-per-class ratio used to develop the algorithms. The normalised cells-per-class ratio calculated using the algorithms in Table 48 can be converted into the relative percentage of one class to another class by treating it as a normal or vulgar fraction in which the numerator gives the percentage of the first class and the denominator the percentage of second class *e.g.* an index of 0.25 indicates a mixture of 20% of the first class to 80% of the second, and an index of 4.0 indicates a mixture of 80% of the first to 20% second.

Table 48. The algorithms for predicting relative contributions of classes for the class identification ratios derived from the 1995 Esthwaite Water reflectance spectra and the integrated sample cell counts.

Phytoplankton classes considered by the class identification ratio (first named : last named)	*Class identification ratio (nm)*	*Slope*	*Intercept*	R^2
Chlorophyte : Cyanophyte	634 : 656	3.65	-3.07	0.438
Chlorophyte : Bacillariophyte	528 : 648	-1.23	1.99	0.596
Cyanophyte : Bacillariophyte	542 : 628	11.93	-12.00	0.396
Cyanophyte : Cryptophyte	610 : 580	-28.51	24.32	0.780
Chlorophyte : Cryptophyte	572 : 658	-10.69	19.05	0.233
Cryptophyte : Bacillariophyte	550 : 516	2.23	-0.99	0.098

These class identification ratios algorithms were then applied to the CASI spectra measured at 1000 m (Table 49). The relative percentage of each class compares well with the cell counts performed upon samples taken at the time of the overflight which showed, in decreasing order of dominance, the presence of Cyanophytes, Cryptophytes and Chlorophytes at approximately the relative concentrations predicted by the class identification ratio algorithms in Table 48.

Table 49. The results of the application of the class identification ratio algorithms to the reflectance spectra measured by the CASI on 13/07/94 at 1000 m.

Class identification ratio classes	*Algorithm result*	*Percentage of each class predicted by class identification ratio algorithm*
Chlorophyte : Cyanophyte	0.39	35% Chlorophyte : 65% Cyanophyte
Chlorophyte : Cryptophyte	0.23	25% Chlorophyte : 75% Cryptophyte
Cyanophyte : Cryptophyte	4.58	75% Cyanophyte : 25% Cryptophyte

The class identification ratios (Table 48) have been successful at predicting the relative concentrations of the various phytoplankton classes in Esthwaite Water. The technique was also successful in predicting the dominance of Cyanophytes over Cryptophytes and Cryptophytes over Chlorophytes from the CASI spectra (Table 49). The most successful class identification ratio for predicting the relative concentrations of the various classes was that of the Chlorophytes to Cryptophytes. The least successful was the Cryptophyte to Bacillariophyte class identification ratio. Interestingly the former

produced the highest F-ratio value and the latter the lowest F-ratio value when the class identification ratios were being developed (Table 20). The Cyanophytes and the Cryptophytes were the two most strongly represented classes throughout 1995 as recorded by the IFE sampling programme. The Bacillariophytes were the most poorly represented class, yet the class identification ratios including the class Bacillariophyceae perform well and appear unaffected by the relative scarcity of cells in this class. This may indicate that the class identification ratios are sensitive to subtle changes in the concentration of each class and that the spectra were above the effective noise floor for these ratios. This noise floor arises when fluxes of reflected radiance are at levels close to the noise floor of the instrument. When these radiance measurements are used in a ratio, any inaccuracies in this measurements are compounded. Under such reflectance conditions ratio-based algorithms can be subject to inaccuracies.

Residual analysis showed that the ability of the Cyanophyte to Cryptophyte class identification ratio to predict the value of the normalised cells-per-class ratio was not correlated with species diversity. A similar conclusion was reached for the less successful Cryptophyte to Bacillariophyte class identification ratio. This suggests that the performance of the class identification ratios to predict the variation in the relative concentrations of two classes is unaffected by the range of different species present. The class identification ratios also appear to be unaffected by variation in the concentration of the other optical parameters.

When the class identification ratios were used to explain variance in the cells-per-class ratios they were more successful if the cell numbers per class were logged. There is considerable justification for performing this form of normalisation because the distribution of cell numbers for each class is log-normal. Taking the natural logarithm reduces the occurrence of outliers in the cells-per-class ratio which arise when a class is either absent or present at very high cell concentrations.

The class identification ratios have performed well in predicting the variation between the concentration of sets of two classes in Esthwaite Water. Algorithms developed from the 1995 reflectance spectra (the spectra measured in the controlled optical environment of the experimental tank) successfully identified the relative concentrations of the three main classes in Esthwaite Water from hyperspectral CASI data. The choice of which of the class identification ratios is most applicable should be made with deliberation of the likelihood of a particular functional group being dominant.

8.3 Application of the canonical discriminant technique to the classification of Esthwaite Water reflectance spectra

The canonical discriminant technique was applied to the Esthwaite Water spectra. Canonical coefficients used were those calculated from the pure culture absorption spectra at the 'analytical' set of wavelengths as developed in Table 27. Spectra from the class Euglenophyceae were excluded from

this calculation because of the relative absence of this class in the Esthwaite Water surface water samples. Canonical coefficients were also calculated for the 'extended analytical' wavelength set for the pure culture $R_{PHY}(\lambda)$ (Table 29). The Esthwaite Water seston absorption spectra, total absorption spectra and sub-surface reflectance spectra were then reduced to the analytical and extended analytical wavelengths respectively and transformed according to the transformation applied to the data set used to calculate the appropriate canonical coefficients. The canonical coefficients were then used to calculate the canonical scores from which classification of the dominant phytoplankton class was made using both the proximity to the nearest class mean and relative distance techniques described in § 7.2.4.1. The success of the classification was measured by comparing the predicted dominant class for each spectra against the dominant class identified in the surface water sample cell counts. Similarly the CASI spectra were reduced to the extended analytical wavelength set, transformed accordingly and the canonical scores calculated using the canonical coefficients from the pure culture $R_{PHY}(\lambda)$ (Table 29). Classification of the dominant class was made and the results similarly compared against the surface sample cell counts.

The Esthwaite Water $a_{SES}(\lambda)$ canonical scores were classified using proximity to the nearest class mean as recorded when the canonical scores of the pure culture $a_{PHY}(\lambda)$ were plotted in canonical space. This classification correctly identified the dominant class in 60% (9 out of 15) of the $a_{SES}(\lambda)$ (Table 50). Misassignments occurred at both high and low levels of phytoplankton diversity, hence a complex $a_{SES}(\lambda)$, comprising a number of $a_{PHY}(\lambda)$, did not reduce the ability of the canonical discriminant analysis to correctly predict the dominant class. Misassignments which included Bacillariophyte spectra were not unexpected because in canonical space the Bacillariophyte class mean was centrally located between the other four classes. The relative distance assignment routine also correctly identified the dominant class of the Esthwaite Water seston absorption spectra with 60% accuracy. As mentioned in § 7.2.4.1 this classification routine also has the ability to identify the sub-dominant class. 33% (5 out of 15) of the $a_{SES}(\lambda)$ had both their dominant and sub-dominant classes correctly identified. Application of the same canonical coefficients to the Esthwaite Water $a_{TOT}(\lambda)$ spectra (§ 5.4.3) reduced the success of the proximity to the nearest mean prediction routine to one spectra out of fifteen.

The canonical scores for the Esthwaite Water $R(\lambda)(-0.01m)$ were computed and the mean class locations of the pure culture $R_{PHY}(\lambda)$ data set (when plotted in canonical space) used to classify the dominant class in each spectra. This classification of the dominant phytoplankton class only achieved a 37% success rate (6 out of 15) (Table 50). In particular the Cyanophytes were misassigned as Bacillariophytes. Attempts to remove the spectral influence of $a_{DYS}(\lambda)$ from the Esthwaite Water reflectance spectra, which was absent from the $R_{PHY}(\lambda)$ data set used to compute the canonical scores, did not improve the prediction success. The relative distance classification routine failed to improve on the success rate.

The canonical scores calculated for the CASI spectra were classified in a similar manner. All were assigned Bacillariophyceae as their dominant class, but the cell counts performed on concurrent samples identified the dominance of Cyanophytes.

Table 50. Summary of the canonical discriminant analysis classifications of the dominant class in the Esthwaite Water spectra. Classes were identified using the nearest class mean classification. Identified classes that concur with the dominant class in the surface water sample cell counts are highlighted.

Sample date	*Dominant class in surface sample cell counts*	*Dominant class identified from the Esthwaite Water* $a_{SES}(\lambda)$	*Dominant class identified from the Esthwaite Water* $a_{TOT}(\lambda)$	*Dominant class identified from the Esthwaite Water* $R(\lambda)(-0.01\ m)$
06/07/94	Cyanophyte	Cyanophyte	Bacillariophyte	Cyanophyte
14/07/94	Cyanophyte	Cyanophyte	Bacillariophyte	Cyanophyte
28/07/94	Cyanophyte	Bacillariophyte	Bacillariophyte	Chlorophyte
01/09/94	Cyanophyte	Cyanophyte	Bacillariophyte	Bacillariophyte
21/09/94	Cyanophyte	Cyanophyte	Bacillariophyte	Bacillariophyte
23/05/95	Chlorophyte	Bacillariophyte	Bacillariophyte	Chlorophyte
06/06/95	Cryptophyte	Bacillariophyte	Bacillariophyte	Bacillariophyte
20/06/95	Cryptophyte	Bacillariophyte	Bacillariophyte	Bacillariophyte
04/07/95	Cyanophyte	Bacillariophyte	Bacillariophyte	Bacillariophyte
25/07/95	Cyanophyte	Cyanophyte	Bacillariophyte	Cyanophyte
01/08/95	Cyanophyte	Bacillariophyte	Bacillariophyte	Bacillariophyte
15/08/95	Cyanophyte	Cyanophyte	Bacillariophyte	Bacillariophyte
12/09/95	Cyanophyte	Cyanophyte	Bacillariophyte	Bacillariophyte
10/10/95	Cyanophyte	Cyanophyte	Bacillariophyte	Cyanophyte
10/04/96	Bacillariophyte	Bacillariophyte	Bacillariophyte	Bacillariophyte

The improved success for the classification of $a_{SES}(\lambda)$ over the $a_{TOT}(\lambda)$ has most probably arisen because of the absorption influences of $a_{DYS}(\lambda)$ and $a_{H2O}(\lambda)$, which, in a Case I water, are the only optically active components along with the $a_{PHY}(\lambda)$. Thus $a_{SES}(\lambda)$ spectra are the closest to the $a_{PHY}(\lambda)$ that were used to the develop of the canonical coefficients.

The relatively poor performance of the canonical discriminant analysis is surprising, particularly as 8 of the 15 Esthwaite Water spectra were acquired when *Anabaena* was the dominant species and *Anabaena* was one of the pure culture species used in the experimental tank. Only three of the Esthwaite Water spectra dominated by *Anabaena* were correctly assigned. The absence of class Cryptophyceae from the pure culture $R_{PHY}(\lambda)$ data set used to compute the canonical coefficients may account for some of misclassifications involving Bacillariophytes because of the relative similarity of pigment composition between Cryptophytes and Bacillariophytes. This may be particularly important because Cryptophytes were present as the sub-dominant class for most of the Esthwaite Water sampling.

Although very successful for the classification of the pure culture $R_{PHY}(\lambda)$, using canonical discriminant analysis as a tool for identifying the dominant phytoplankton class in the Esthwaite

Water optical spectra has not been particularly successful. The application of canonical discriminant analysis to the classification of remote sensing data is hampered by the reliance placed upon the class distributions of the data sets used to compute the canonical scores and ascertain the location of the class means within canonical space for the purposes of prediction. This is the most likely explanation of why this technique was only partially successful when applied to measurements made in the natural environment. As an indication of how dependent this technique is on the 'training' data set it is interesting to note that changing the canonical coefficient at 400 nm for the first canonical function from its value of 2125.8 (Table 29) to -280.0 improves the prediction of the dominant class from Esthwaite Water reflectance spectra from 37% to 68%. This optimisation for an arbitrarily selected coefficient shows how erratic this technique can be.

8.4 Application of derivative analysis to classification of Esthwaite Water reflectance spectra

The derivative analysis technique was then used to classify the dominant and sub-dominant classes from the Esthwaite Water reflectance spectra. The 4^{th} derivative {4,4,29,1} of the Esthwaite Water $R(\lambda)(-0.01m)$ were calculated and the troughs identified at a resolution of 5 nm. The locations of these features were compared to the characteristic derivative finger-prints identified for each class as described in § 7.3.4.2 and a dominant and sub-dominant class assigned accordingly.

80% of the Esthwaite Water $R(\lambda)(-0.01$ m) (12 out of 15) were correctly assigned a dominant class as confirmed by cell counts performed upon surface water samples taken contemporaneously with the spectral signatures (Table 51). Three of the misclassifications involved identifying the class Chlorophyceae as a Cryptophyceae or *vice-versa.* The dominance of the Chlorophytes on 23/05/95 was not identified, nor was the dominance of the Bacillariophytes on 10/04/96. Both these spectra were misclassified as being dominated by the class Cyanophyceae which was poorly represented on these sample dates. Interestingly the misclassification of a sample was more likely to occur when the species diversity index was low *i.e.* when fewer species, hence when there were fewer conflicting species reflectance spectra in the combined spectral signature. The accuracy of prediction was reduced 60% when both the dominant and sub-dominant classes were considered.

Table 51. The actual dominant and sub-dominant classes and those predicted using derivative analysis of the Esthwaite Water R(λ). Correct predictions are highlighted.

Sample date	*Actual dominant class in surface sample cell counts*	*Actual sub-dominant class in surface sample cell counts*	*Predicted dominant class*	*Predicted sub-dominant class*
06/07/94	Cyanophyte	Chlorophyte	Chlorophyte	Cryptophyte
14/07/94	Cyanophyte	Chlorophyte	Cyanophyte	Chlorophyte
28/07/94	Cyanophyte	Chlorophyte	Cyanophyte	Chlorophyte
01/09/94	Cyanophyte	Cryptophyte	Cyanophyte	Cryptophyte
21/09/94	Cyanophyte	Cryptophyte	Cyanophyte	Cryptophyte
23/05/95	Chlorophyte	Cryptophyte	Cyanophyte	Bacillariophyte
06/06/95	Cryptophyte	Bacillariophyte	Cryptophyte	Bacillariophyte
20/06/95	Cryptophyte	Bacillariophyte	Cryptophyte	Cyanophyte
04/07/95	Cyanophyte	Cryptophyte	Cyanophyte	Cryptophyte
25/07/95	Cyanophyte	Cryptophyte	Cyanophyte	Chlorophyte
01/08/95	Cyanophyte	Cryptophyte	Cyanophyte	Chlorophyte
15/08/95	Cyanophyte	Cryptophyte	Cyanophyte	Cryptophyte
12/09/95	Cyanophyte	Cryptophyte	Cyanophyte	Cryptophyte
10/10/95	Cyanophyte	Cryptophyte	Cyanophyte	Cryptophyte
10/04/96	Bacillariophyte	Cryptophyte	Cyanophyte	Bacillariophyte

The 4th derivative {4,4,29,1} of the CASI spectra were then calculated and the trough features identified at a resolution of 5 nm. These 4th derivative finger-prints were then assigned a dominant and sub-dominant class which was compared to the actual classes identified in the contemporaneous cell counts. The presence of the 630 nm trough feature lead to the prediction of Cyanophytes as the dominant class for all the sampled spectra (Table 52). The incorrect prediction of the class Chlorophyceae as the sub-dominant class arose due to the presence of the 490 nm trough feature which is typically associated with absorption by xanthinols. With the exception of the 615 nm feature (R-PCC) identified in the spectra from sites 5 and 6 there is no indication of the presence of Cryptophytes which were found to be sub-dominant in the cell counts. According to the cell counts Cyanophytes were very dominant so it could be expected that the Cyanophyte signal masked the underlying signal from the sub-dominant classes. However it may also be expected that the derivative spectra should exhibit more of the features found in the Cyanophyte 4th derivative spectra (Table 41). The absence of more Cyanophyte marker features, in particular the 555 nm feature, in the CASI 4th derivative spectra is noteworthy. The absence of these features in this data set could be ascribed to the fact that the CASI spectral profile were recorded at 7 nm intervals and splined to 2 nm to comply with the above analysis. Alternatively the atmospheric correction described in § 5.5.2 could have been responsible for some of the spectral features present in both the surface corrected $Lu_a(\lambda)$ or the final R(λ)(-0.01 m).

Table 52. The 4th derivative {4,4,29,1} Esthwaite Water CASI spectra trough locations with predicted dominant and sub-dominant classes according to the derivative analysis and the actual classes confirmed by surface water sample cell counts. Correctly identified classes have been highlighted.

Trough feature number	*Site 1*	*Site 2*	*Site 3*	*Site 4*	*Site 5*	*Site 6*
1	440	440	440	440	440	450
2	490	490	490	485	490	480
3	530	530	530	530	530	475
4	580	580	580	580	580	490
5	610	610	610	610	615	530
6	630	630	630	630	630	580
7	680	680	680	680	680	615
8	-	-	-	-	-	630
9	-	-	-	-	-	680
Actual dominant class	Cyanophyte	Cyanophyte	Cyanophyte	Cyanophyte	Cyanophyte	Cyanophyte
Actual sub-dominant class	Cryptophyte	Cryptophyte	Cryptophyte	Cryptophyte	Cryptophyte	Cryptophyte
Predicted dominant class	Cyanophyte	Cyanophyte	Cyanophyte	Cyanophyte	Cyanophyte	Cyanophyte
Predicted sub-dominant class	Chlorophyte	Chlorophyte	Chlorophyte	Chlorophyte	Chlorophyte	Chlorophyte

The application of derivative analysis to the classification of phytoplankton class in the reflectance spectra from this natural Case I water body has been successful, showing that the technique transposes easily from the laboratory to the natural environment. Both the dominant and sub-dominant phytoplankton classes were identified for most of the Esthwaite Water reflectance spectra. Neither the subjective nature of the derivative feature location (§ 7.3.5) nor the 5 nm resolution appears to have hampered the technique. Furthermore it is possible that the derivative class identification technique will be applicable to all Case I water bodies without the need for modification.

8.5 Application of neural networks to the classification of Esthwaite Water reflectance spectra

The neural network phytoplankton class identification technique was then used to classify the dominant class from the Esthwaite Water reflectance spectra. The trough locations of the 4th derivative Esthwaite Water $R(\lambda)(-0.01\ m)$ were identified at a resolution of 5 nm and the neural network trained using the pure culture $a_{PHY}(\lambda)$ (§ 7.4.3). This trained network was used to predict the dominant and sub-dominant classes. The results of the prediction were compared against the cell counts.

The neural network successfully and positively identified 67% (10 of 15) of the Esthwaite Water reflectance spectra as confirmed by cell counts from surface water samples (Table 53). If a classification by the network is deemed successful if either the dominant or sub-dominant class was correctly identified then the network's success improved to 80% for the surface sample cell counts.

It can be seen from Table 53 that the neural network has positively identified each spectra as belonging to one definitive class and has applied very low scores to the other classes. The network predictions have not provided any information pertaining to the relative contributions of each class. Misclassification of spectra tended to be for those spectra where a dominant class was only weakly dominant. For such spectral samples it is possible that a mixture of weak spectral signals would confuse the recognition routines of the network, hence lead to a misclassification.

Table 53. The results of using the neural network trained on pure culture $a_{PHY}(\lambda)$ for the classification of the Esthwaite Water $R(\lambda)(-0.01\ m)$. The network score for each of the five classes are given, with the highest score in bold. Correctly classified spectra are highlighted. The dominant classes in the surface sample cell counts are also given. Samples where class dominance was relatively weak are marked with a *.

Sample date	*Network scores*					*Network classified result*	*Actual dominant class in surface sample*
	Chloro-phyte	*Cyano-phyte*	*Bacillario-phyte*	*Crypto-phyte*	*Eugleno-phyte*		
06/07/94	**0.82**	0.22	0.02	0.08	0.01	Chlorophyte	Cyanophyte
14/07/94	**0.84**	0.13	0.02	0.11	0.06	Chlorophyte	Cyanophyte
28/07/94	0.04	**0.98**	0.02	0.05	0.02	Cyanophyte	Cyanophyte
01/09/94	0.01	**0.82**	0.42	0.08	0.06	Cyanophyte	Cyanophyte
21/09/94	0.06	**0.96**	0.03	0.08	0.03	Cyanophyte	Cyanophyte
23/05/95	0.03	**0.82**	0.09	0.14	0.06	Cyanophyte	Chlorophyte*
06/06/95	0.05	0.18	0.03	**0.64**	0.02	Cryptophyte	Cryptophyte*
20/06/95	0.00	**0.96**	0.18	0.04	0.02	Cyanophyte	Cryptophyte*
04/07/95	0.12	0.26	**0.72**	0.00	0.02	Bacillariophyte	Cyanophyte*
25/07/95	0.02	**0.99**	0.01	0.03	0.08	Cyanophyte	Cyanophyte
01/08/95	0.02	**0.93**	0.01	0.11	0.01	Cyanophyte	Cyanophyte
15/08/95	0.02	**0.98**	0.02	0.13	0.04	Cyanophyte	Cyanophyte
12/09/95	0.01	**0.69**	0.03	0.39	0.01	Cyanophyte	Cyanophyte
10/10/95	0.01	**0.99**	0.02	0.07	0.04	Cyanophyte	Cyanophyte*
10/04/96	0.01	0.04	**0.84**	0.03	0.00	Bacillariophyte	Bacillariophyte

The 4th derivative trough locations of the reflectance spectra obtained at altitude using the CASI (Table 52) were then classified using the neural network as trained on the pure culture absorption spectra. The network successfully identified the dominant class in four of the six spectral profiles. None of the successful identifications were definitive in that the network gave the Cyanophytes (the dominant class present in cell counts performed on samples concurrently with the overflight) scores of 0.74 to 0.98 but also gave scores of 0.6 to the Bacillariophytes. Indeed the two misclassified spectral profiles suggested that Bacillariophytes were the dominant class. The cell counts indicate that the sub-dominant class was actually Cryptophytes which consistently scored poorly in the classification.

Neural networks have therefore successfully identified the dominant class in both surface and altitude reflectance spectra. As with the application of the derivative analysis technique the neural network application transferred without need for modification. Consequently it should prove applicable to the problem of class identification from any Case I water body.

8.6 Conclusions from the application of the class identification routines to the Esthwaite Water reflectance spectra

The various techniques for the identification of dominant phytoplankton class through remote sensing have been tested on reflectance spectra from a natural water body. They have achieved differing levels of success as summarised in Table 54.

Table 54. Comparison of the success of the various class identification techniques at identifying the dominant class in the sub-surface and altitude Esthwaite Water reflectance spectra.

Classification technique	*% of the dominant class correctly identified from the 13 R(λ)(-0.01 m)*	*% of the dominant class correctly identified in the 6 R(λ)(1000 m)*
Number of peaks	82 %	100 %
Green peak location	87 %	100 %
Class ratio	N/A	N/A
Canonical discriminant	37 %	0 %
Derivatives	80 %	100 %
Neural networks	67 %	66 %

The routine based on the number of peaks exhibited by each spectra relied upon the distinctive three peaked reflectance spectra that is characteristic of the Cyanophytes. This technique can only distinguish between Cyanophytes and non-Cyanophytes so would not be applicable to those lakes which do not support Cyanophyte blooms. Thus the ability to identify the presence or absence of Cyanophytes is restrictive given the range of possible freshwater phytoplankton classes (Christiansen 1962). However, the as Cyanophytes are the class most likely to cause water quality problems when they attain bloom concentrations, the ability to identify Cyanophytes is useful.

The location of the green reflectance peak was shown to vary by class in the pure culture $R_{PHY}(\lambda)$ due to variation of the location of the $a_{PHY}(\lambda)$ minimum at *ca.* 550 nm (Table 9). The location of the green reflectance peak therefore lends itself to being used as a classification routine. However, this particular routine required a wavelength shift adjustment of 20 nm before it successfully identified the dominant class in 87% of the Esthwaite Water spectra. Therefore this routine requires further validation across a number of freshwater lakes to establish whether this adjustment in location of the reflectance peak was site specific to Esthwaite Water or was necessary because of the optical characteristics of the experimental tank.

The correspondence between the class reflectance ratio for the Cyanophytes and Cryptophytes and the actual ratio of the number of cells from these two classes (Figure 79) was high. The ratio successfully identified those samples dominated by Cryptophytes those samples dominated by Cryptophytes. However, because these class identification ratios were restricted to two classes only the application of the class ratio routine could lead to an inaccurate classification if the wrong class ratio is used. Another drawback to the general application of ratio-based algorithms is their susceptibility to the

influence of signal noise. Using data close to the noise floor of the ratios can give misleading results (§ 8.2). The spectral signal from Esthwaite Water was typically strong because of the high phytoplankton biomass, however spectra from oligotrophic waters will reflect lower volumes of radiance hence will be susceptible to noise. Application of this ratio based classification routine to the reflectance spectra from such lakes should be made with caution.

The canonical discriminant analysis which proved very successful at classifying the pure culture $a_{PHY}(\lambda)$ and $R_{PHY}(\lambda)$. However this technique failed to classify the Esthwaite Water spectra because the classification procedure was reliant upon the distributions of the data set from which the canonical correlations were calculated. This classification routine will always require transformation of the unknown spectra into the same format as those spectra used to compute the canonical coefficients. It is for this reason that the canonical coefficients computed for the more extensive pure culture $a_{PHY}(\lambda)$ could not be used in the classification of pure culture $R_{PHY}(\lambda)(-0.01\ m)$. Similarly the canonical coefficients calculated from the pure culture $R_{PHY}(\lambda)(-0.01\ m)$ measured in the experimental tank could not be used to classify the natural $R(\lambda)(-0.01\ m)$ measured from Esthwaite Water .

The derivative analysis routine proved to be the most successful, correctly classifying 80% of the dominant 60% of the sub-dominant classes from the Esthwaite Water reflectance spectra. No modifications were required from the technique developed from the $a_{PHY}(\lambda)$ to perform this classification, consequently the technique lends itself particularly well to remote sensing. The neural network application drew upon the derivatives technique and used the presence or absence of peak features to essentially recognise the 'bar-code' of the spectra. The neural network technique was less successful than the derivative analysis technique in the classification of the Esthwaite Water $R(\lambda)(-0.01\ m)$ and an explanation for the classification success was hidden within the neural network.

Those techniques which utilised more of the information contained in the hyperspectral reflectance spectra (*i.e.* the derivative based routines, the peak location and peak determination routines) were more adept at the classification. Those routines using a limited number of spectral bands as inputs (*i.e.* the class identification ratios and the canonical discriminant analysis) were less successful for classifying spectra from Esthwaite Water.

The characteristics of a classification technique that can satisfy the aim of this study is one which can not only classify spectra obtained in the laboratory, but one which can also classify remotely sensed spectra obtained at altitude from a natural water body. Only a technique which can perform this transition from laboratory to field will further remote sensing as a tool for the water manager. Of the six techniques tested upon the Esthwaite Water reflectance spectra the 4^{th} derivative analysis technique has made the most successful transition from classifying pure culture $a_{PHY}(\lambda)$, through pure culture $R_{PHY}(\lambda)$, to field and airborne spectra from the study lake. It has been shown that this

technique enhances the absorption characteristics of the class marker pigments as manifest in the reflectance spectra. The derivative technique required no adjustment to classify the Esthwaite Water spectra from the technique developed from the $a_{PHY}(\lambda)$. Furthermore the success of the derivative analysis technique was not reduced by the presence of non-phytoplanktonic signals in the Esthwaite Water $R(\lambda)(-0.01m)$. This includes those signals from the other optically active water quality constituents (tripton, DYS and water itself), from the spectral error features introduced by performing measurements in the field and from the spectral errors introduced during the conversion of the upwelling above surface radiance spectra into below surface irradiance reflectance spectra.

Chapter Nine Conclusions

The aim of this research was to improve remote sensing as a tool for water managers by developing techniques for the quantification and qualification of freshwater phytoplankton from remotely sensed data. There were two objectives to this research: assessing the applicability of CHL*a* retrieval algorithms designed for different phytoplankton classes and assessing the ability of remote sensing to detect the presence of different freshwater phytoplankton classes. Both these objectives have been realised. Of the classification routines developed, a 4^{th} derivative analysis technique was shown to be the most promising for classifying freshwater phytoplankton from remotely sensed reflectance spectra. CHL*a* retrieval algorithms were developed and compared against published universal algorithms. It was shown that empirical algorithms are the most accurate for remotely sensing CHL*a* concentration on an individual lake basis and that the accuracy of CHL*a* retrieval could not be significantly improved when class specific CHL*a* algorithms were employed. With the fulfilment of these objectives this research has achieved the research aim of improving the applicability of remote sensing as a tool for freshwater management.

The benefits of this aim are that improved monitoring of both the spatial distribution of CHL*a* in a lake and the spatial variation in the composition of the phytoplankton crop should allow water managers to adopt more efficient and effective remediation measures. In freshwater systems the nutrient most likely to limit the generation of a phytoplankton bloom (given the presence of a base population) is phosphorus. Treatment of nuisance freshwater blooms is generally based around the removal of sources of phosphorus. Among other methods phosphorus can be controlled by: the removal of phosphate from point and diffuse inputs (sewage works and farmland); removal of phosphate from the water column and removal of phosphate rich sediments from the lake bed. Alternatively the onset of lake stratification can be prevented by forced circulation or various biomanipulation techniques can be employed (Shapiro 1993). The more information that can be given to water managers pertaining to the phytoplankton bloom (its dominant class, its composition, the bloom concentration and the bloom's physical extent) the more likely that the appropriate control measure will be selected.

9.1 Summary of findings

The various freshwater phytoplankton classes exhibit different inherent optical characteristics because they contain different combinations of intracellular accessory photosynthetic pigments. The presence of certain accessory pigments is restricted to individual phytoplankton classes, hence these class marker pigments can be used to indicate the presence of particular classes. The pigments absorb at discrete wavelengths within the visible electromagnetic spectrum which imparts colour to the phytoplankton cells. The variations in pigment composition and therefore phytoplankton absorption

are a dominant influence on the apparent optical property phytoplankton reflectance. Phytoplankton reflectance spectra can measured by airborne remote sensing instruments in both low and high spectral resolution. Theoretically it should therefore be possible to use remote sensing to identify the presence of class pigments and therefore classify the dominant class of the target bloom.

The optical properties of those pigments likely to be present in freshwater phytoplankton were introduced and the pigments most likely to serve as class indicative marker pigments were identified. Measurements of the absorption from pure cultures of phytoplankton from different classes were presented. Although the absorption by CHL*a* was dominant in every class, the absorption by the accessory pigments were identifiable in these spectra. The scattering spectra for these species were approximated and their angular scattering distributions measured. Spectral variations in scattering, and variations in scattering propensity, were found to be independent of class type. As such variations in reflectance introduced by phytoplankton scattering could not be used to assist in the classification.

Measurements of reflectance from pure cultures of phytoplankton were made in an experimental tank constructed for this study. The magnitude of the reflected flux varied directly with the increased scattering, occurring indifferently across the visible spectrum and linked directly to phytoplankton concentration. The spectral features, that allowed the reflectance spectra from the three phytoplankton classes to be distinguished, were shown to be related to pigment composition. In particular these features were shown to be the result of the absorption by the accessory pigments. It was noted that between class differences in spectral form that could be attributed to pigment induced variations were less obvious in the phytoplankton reflectance spectra than in the phytoplankton absorption spectra.

Esthwaite Water, a natural water body in the English Lake District, was shown to be a Case I water body with its optical properties intrinsically linked to the composition and concentration of the phytoplankton crop. Time-series Esthwaite Water absorption and reflectance spectra covering two growing seasons were presented. This undertaking resulted in a database of both absorption and reflectance spectra with supporting water quality analysis which has considerable potential for improving the understanding of how the optical quality of a lake changes over the year. In particular it has proved useful for the testing of both the CHL*a* retrieval algorithms and the class identification routines developed in this study.

CHL*a* retrieval algorithms were then developed from the pure culture and the Esthwaite Water reflectance spectra. Ratio-based algorithms were developed from the pure culture reflectance spectra. Minor improvements in the accuracy of CHL*a* measurement could be achieved if class specific algorithms were used. The algorithms were then tested on the sub-surface and aerial Esthwaite Water reflectance spectra. The performance of these algorithms was compared against three universal algorithms presented in the literature. The empirical algorithm developed from the Esthwaite Water

spectra was considerably more successful at predicting the CHL*a* concentrations than were the universal algorithms. Tailoring the CHL*a* retrieval algorithm to class type did not improve the accuracy of CHL*a* prediction for the natural water body.

Four class identification routines were developed. These were: class ratios; canonical discriminant analysis; 4th derivative analysis; and neural networks. These were developed on pure culture absorption spectra and applied to the pure culture reflectance spectra. These routines were tested on both the Esthwaite Water field reflectance spectra and airborne spectra. Although all the routines were successful at identifying the class type of the pure culture phytoplankton reflectance spectra in the controlled conditions of the experimental tank, the class ratio and canonical correlation routines were unsuccessful when applied to the reflectance spectra from the natural water body.

9.2 Conclusions for the chlorophyll-*a* retrieval algorithms

One of the two objectives of this study was to ascertain how universal remote sensing CHL*a* retrieval algorithms were affected by variations in the composition of the target bloom. Supplementary to this was a comparison of the empirically derived CHL*a* retrieval algorithms against universal algorithms presented in the literature.

Both the empirically derived ratio based algorithms Equation 6.3.1 and Equation 6.2.2.1 used the far-red and near-IR features of the phytoplankton reflectance spectra. For mesotrophic and eutrophic Case I freshwater lakes remote sensing algorithms using a ratio of these observational wavelengths will provide the best means for predicting CHL*a* concentration. As was discussed in § 6.2.2 the far-red reflectance trough is the location of the *ca.* 678 nm CHL*a* absorption peak. This occurs in relative isolation from the absorption signals of both accessory pigments and other optical active water quality parameters. The strength of CHL*a* absorption at 678 nm is directly related to the phytoplankton concentration. Absorption by CHL*a* at this wavelength reduces the amount of light available for scattering and therefore generates a reflectance trough at *ca.* 678 nm. The near-IR reflectance peak at *ca.* 710 nm arises from a window in the combined absorption spectra for phytoplankton and water. Hence increases in phytoplankton concentration results only in increased scattering, reflectance is not moderated by absorption. Therefore increases in phytoplankton concentration induces two different optical effects which have contrasting effects upon reflectance at these wavelengths. These effects are utilised by these ratio based CHL*a* retrieval algorithms because the ratio's product correlates strongly with CHL*a* concentration (Figure 45 and Figure 49).

Algorithms developed from the pure culture reflectance spectra showed that class specific CHL*a* algorithms (Table 15) were more accurate than the algorithm developed from the combined set of pure culture phytoplankton reflectance spectra (Equation 6.2.2.1). However, due to the optimised optical conditions of the experimental tank environment which facilitated accurate reflectance measurements,

Equation 6.2.2.1 explained a very high percentage of the variance in measured CHL*a* concentration (96%). Any increases in accuracy over this value that were achieved by the class tailored algorithms were unlikely to be any more informative. Hence for the pure class phytoplankton reflectance spectra measured in the controlled conditions of the experimental tank class tailored CHL*a* retrieval algorithms were only slightly more accurate than the generic CHL*a* retrieval algorithm developed for all the classes.

The most accurate algorithm was the empirical lake specific algorithm developed from the Esthwaite Water field spectra CHL*a* (μgL^{-1}) = 103.5 ($R_{(722)}$(-0.01 m) : $R_{(684)}$ (-0.01 m)) - 62.8 (Equation 6.3.1). When the Equation 6.2.2.1 was applied to the reflectance spectra from the natural phytoplankton crop on Esthwaite Water the percentage of explained variance was 63% (Table 19). The class tailored algorithms (Table 15) did not enhance the accuracy of CHL*a* retrieval.

The class tailored algorithms in Table 15 exhibited subtle variations in the location of the observation wavelength for the numerator and denominator wavelengths centering upon 697 ± 6 nm : 673 ± 5 nm. These variations exist because the location of the far-red CHL*a* absorption feature varies for each class vary. As was highlighted by Table 10 the locations of the far-red reflectance troughs were found to be very similar to the locations of the denominator wavelengths for the class tailored CHL*a* algorithms. It would therefore appear that the phytoplankton classes used in this study do have different forms of CHL*a*. These CHL*a* forms exhibit subtly different absorption features which could be used to develop more sophisticated and more successful class tailored CHL*a* retrieval algorithms. Further study is necessary to establish whether these shifts in the CHL*a* absorption maxima occur because of differences in the chemical structure of the CHL*a* in each class, or because identical forms of CHL*a* are existing in the presence of a range of different accessory pigments and within a range of differently formed cellular structures.

The performance of Equations 6.3.1, 6.2.2.1 and the class tailored algorithms in Table 15 were compared against some universal algorithms presented in the literature using Esthwaite Water reflectance spectra as a test data set. The results of this comparison exercise (Table 19) has clearly shown that the most appropriate CHL*a* retrieval algorithm for this Case I water body was not the universal algorithms presented in the literature, nor the class tailored algorithms, but the empirical algorithm developed from the Esthwaite Water field measurements.

Universal CHL*a* retrieval algorithms have to perform accurately regardless of the particular optical quality of the target water body. Lake optics are dependant upon the nature of the incident radiation and the composition of the various optically active water quality parameters. Although the ratio wavelengths for Equation 6.3.1 are at similar locations to those used in the universal algorithm of Dekker (1993), the algorithm parameters differ. It is the ability to tailor these parameters to individual lakes that gives the empirical algorithm, based on lake specific ground truth data, the improved

accuracy over universal algorithms. This was also true when the parameters of Equation 6.2.2.1 were optimised for the various target phytoplankton classes (Table 14). Subtle changes in the algorithm parameters achieved improved accuracy of CHL*a* prediction. Although the goal of the universal algorithm is a necessary one if remote sensing is to prove a cost effective tool for the water manager, universal algorithms are not yet sufficiently developed to provide the accuracy of CHL*a* retrieval that can be realised from empirical algorithms. Consequently, if accurate CHL*a* measurements are to be achieved through remote sensing, there will always be a need for ground truth measurements to facilitate the development of empirical algorithms. The R-matrix method would appear to be a particularly succinct method of establishing the most appropriate empirical algorithm from a set of ground truth reflectance spectra.

9.3 Conclusions for the class identification routines

To be a cost effective tool for water managers remote sensing must provide more information than solely the concentration of a phytoplankton bloom. Effective remediation requires knowledge of the bloom composition. This requirement provided the second objective for this study, focusing upon the development of remote sensing classification routines. Various classification routines have been developed. From their performance at classifying the dominant phytoplankton class from a natural water's reflectance spectra it was apparent that the 4^{th} derivative analysis technique was the most appropriate for implementation as an operational water management tool. A simple routine based upon the presence of a reflectance peak at 665 nm also showed potential to be a simple technique for confirming the presence of nuisance Cyanophyte blooms.

The development of the classification routines drew heavily on the assumption that phytoplankton reflectance is the inverse of phytoplankton absorption and that phytoplankton scattering has little influence on the spectral shape of phytoplankton reflectance. Measurements conducted during this study and the scattering efficiency spectra presented by Bricaud and Morel (1986) and Bricaud *et al.* (1988) show scattering for phytoplankton to be typically non-trending and to exhibit only minor spectral features in the vicinity of the strong CHL*a* absorption features. This assumption was required because phytoplankton scattering typically varies with cellular morphology which varies on a species, not class, level. Differences in phytoplankton scattering will result in variation in the magnitude of phytoplankton reflectance, but in the absence of absorption this reflectance will be flat and featureless throughout the visible spectrum. Such a spectrum could not be used to assist the classification. Phytoplankton absorption is primarily a function of pigment concentration and composition. The latter varies on a class level. Thus variations in phytoplankton absorption, with the concomitant variation in phytoplankton reflectance, allows the successful classification of phytoplankton class to be made.

Phytoplankton contain a variety of photosynthetic pigments including CHLa (which is ubiquitous across all classes) and accessory pigments. Because certain accessory pigments are unique to particular classes they can be used as class markers. The distinctive absorption properties of these pigments can be distinguished in the phytoplankton reflectance spectra leading to successful classification. However, individual pigment features can be masked by more strongly absorbing pigments making the task of identifying specific class marker pigments difficult. The class marker pigments used in this study are shown in Table 55.

Table 55. Class marker pigments used in this study with their absorption wavelengths.

Phytoplankton class	*Class marker pigments*	*Absorption wavelength (nm)*
Euglenophytes	DDX	490
Bacillariophytes	FCX, CHL*c*	530, 465
Cryptophytes	R-PCE	575
Cyanophytes	C-PCC, MYX	630, 500
Chlorophytes	CHL*b*	656

The classification routines were initially developed on the pure culture phytoplankton absorption spectra where the between-class differences resulting from differences in pigment composition were most obvious. Once shown to work on the pure culture phytoplankton absorption spectra the routines were transposed to the pure culture phytoplankton reflectance spectra. This approach, moving from the inherent to the apparent optical properties of phytoplankton, meant that the routines could be developed in the relative absence of signal noise. Furthermore, because the pure culture phytoplankton absorption spectra were easier to measure than the pure culture phytoplankton reflectance spectra a larger variety of species from each class could be measured. Finally, by nature of their pseudo-analytical basis, the classification routines thus developed should be applicable not just to the phytoplankton reflectance spectra measured in the experimental tank but also to reflectance spectra from natural water bodies.

A 4^{th} derivative analysis technique proved to be very successful at classifying the dominant phytoplankton class from the Esthwaite Water reflectance spectra. Passing a 4^{th} derivative through hyperspectral reflectance spectra highlighted the subtle spectral of the class marker pigments which were otherwise masked by the spectral properties of the other optically active components. The presence of these pigments could then be more readily identified and a classification made. Applying this technique to the 4th derivative pure culture absorption spectra highlighted a small set of features with distinctive wavelength locations (Table 39). These locations were found to be the centre of absorption for the class marker pigments identified in the literature (Table 1).

The derivative technique utilised the hyperspectral reflectance spectra in a more comprehensive way than the other routines. It required no modification to successfully identify the dominant class in pure culture absorption and reflectance spectra, in Esthwaite Water $R(\lambda)(-0.01\ m)$ and Esthwaite Water

R(λ)(+1000 m). The technique can be readily applied to hyperspectral remote sensing data. It is fundamentally more flexible than the other hyperspectral based classification techniques. Not only does it provide a classification of the dominant class but also of the sub-dominant class. Furthermore it should be easily adapted to include other phytoplankton class types not considered during this study if the location of their class marker pigments is known and does not occur within the close vicinity of the other accessory pigments.

Given access to the CASI the 4th derivative technique would be applied in the following manner. It would be necessary to make two passes of the target lake. During the first pass the CASI would be operated in spectral mode. Recorded upwelling radiance would be processed into sub-surface reflectance on board the plane in real time. The reflectance would be converted into the 4th derivative and the trough features identified. The classification of the dominant class of the target bloom would be made on the basis of the presence of features occurring at the wavelengths in Table 39. Having made the classification the lake would be re-flown with the CASI operating in spatial mode. The pre-selection of the location and bandwidth of the 15 spatial bands would be made by reference to an expert system. This expert system would select the appropriate bands on the basis of the dominant class, thereby optimising the information retrieval possibilities from the flight. For instance, following the identification of a Cyanophyte bloom, a band centred on 630 nm would be appropriate for mapping PCC distributions (Dekker 1993). Optimising the information retrieval of a remote sensing campaign will offset the cost of such campaigns and cost is currently prohibitive to the regular application of remote sensing as a tool for freshwater management.

The classification routine involving assessment of the number of peaks in the hyperspectral reflectance spectra also proved effective. This routine utilised the third reflectance peak generated by the presence of the Cyanophyte marker pigment C-PCC. However it is restricted to identifying the presence or absence of Cyanophytes. This drawback is not overly restrictive because most of the problematic phytoplankton in freshwaters are Cyanophytes, so the ability to confirm the presence of a Cyanophyte bloom may be sufficient to meet the needs of some freshwater managers.

The other classification routines developed in Chapter Seven were both less successful at classifying the Esthwaite Water reflectance spectra than the derivative analysis and they showed less potential for development into operational classification routines. This does not preclude their inclusion in this study because although they failed in the transition from the laboratory to the natural environment, they have all proved successful at classifying pure culture absorption spectra. As such they are useful analytical tools. Of these routines the most disappointing was the canonical discriminant analysis. This technique was very useful for reducing hyperspectral data to a handful of observations whilst maintain enough information to successfully classify the pure culture absorption and reflectance spectra. However, as a parametric technique, it was reliant upon the statistical distributions of the data set used to calculate the canonical correlations. It was therefore unable to successfully classify spectra

not obtained under the same optical conditions as this data set. This notwithstanding as a means for highlighting those wavelengths rich in information pertaining to class types the canonical correlation analysis was very useful. Using the same reasoning the class ratio technique proved useful in the development stage as it confirmed the hypothesis that the spectral features of marker pigments could be used to discriminate between phytoplankton classes. It also proved to be disappointing in the application stage. The neural network technique trained upon the presence of peaks in the 4^{th} derivative absorption spectra was moderately successful at classifying the Esthwaite Water reflectance spectra. As the neural network developed its own classification criteria these were hidden within the network so it was not possible to assess whether the network was utilising derivative features caused by the absorption properties of class indicative marker pigments. The lack of understanding of the classification criteria is the most obvious drawback to the neural network classification routine.

9.4 Final considerations

Since the late 1980's toxic Cyanophyte blooms have been reported to be occurring in many freshwater lakes and reservoirs in the UK (NRA Report 1990). These blooms have been shown to occur in patches, accumulating at the water's surface under the influence of local bathymetry and winds (George 1993). Airborne remote sensing techniques provide the ideal means for recording the formation and dissolution of these patches. Optimising the information retrieval through the application of the findings in this study will enhance remote sensing as a tool for both limnologists and water managers.

The CHL*a* retrieval algorithms developed for the Case I waters of the experimental tank and Esthwaite Water used reflectance ratios of the near-IR to the far-red. The selected denominator wavelength is where the absorption by CHL*a* occurs unobscured by the absorption from accessory pigments. At the numerator wavelength there is no pigment related absorption by phytoplankton with spectral variation due entirely to scattering. The classification routines used observations from the blue to mid-red wavelengths, on the whole giving little significance those wavelengths where the absorption by the accessory pigments is obscured by strong CHL*a* signals. This is because CHL*a* is ubiquitous, it provides limited information pertaining to class type and when present is typically the dominant phytoplankton absorption characteristic within the visible spectrum. The reflectance spectrum can effectively be split into two regions: one where accessory pigment absorption occurs; and one where CHL*a* related absorption occurs (Figure 80). There is little overlap, with classification routines utilising the accessory pigment features and the CHL*a* retrieval algorithms avoiding the accessory pigment absorption features. This generalisation holds for Case I mesotrophic and eutrophic waters where the ratio based CHL*a* algorithm was shown to be most appropriate. In Case I oligotrophic waters the most appropriate CHL*a* retrieval algorithm will be a green to blue ratio (Equation 6.2.2.3).

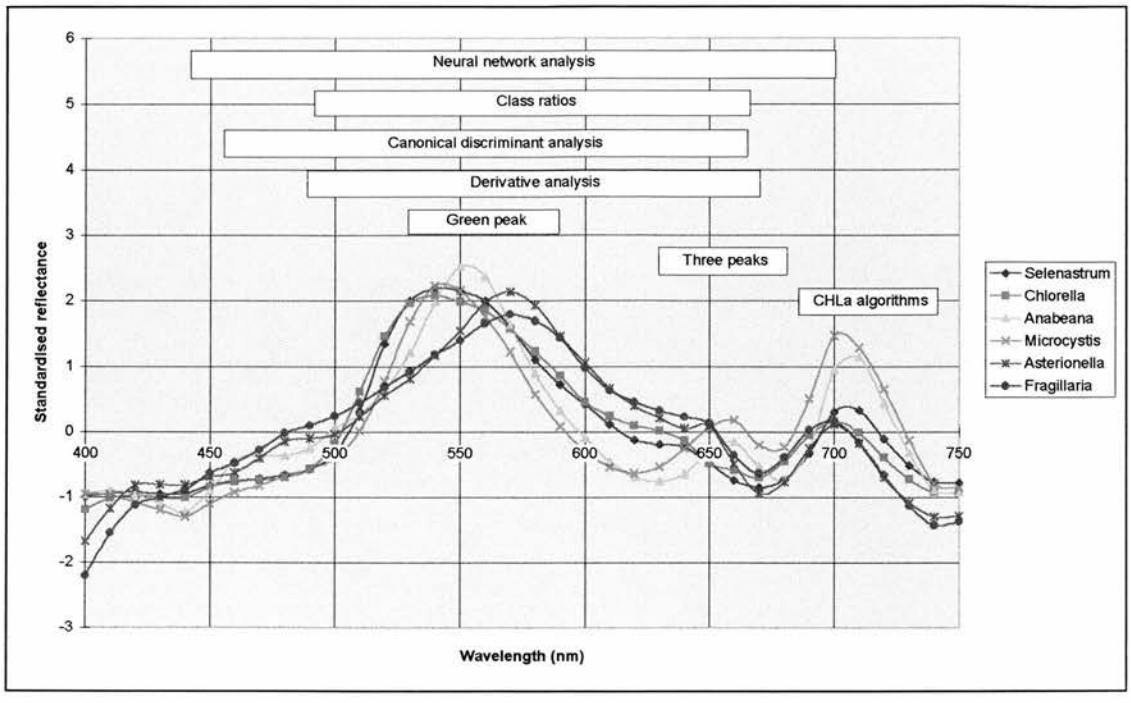

Figure 80. The input observational wavelengths required for the classification routines and the CHL*a* algorithms. These have been superimposed above the standardised reflectance spectra for the pure cultures measured in the experimental tank shown in Figure 30.

The aim of the universal CHL*a* retrieval algorithm is to circumvent the requirements for the supplementary field measurements necessary for the calculation of empirical algorithms for individual lakes. Performing these field measurements adds to the cost of remote sensing campaigns and decreases the cost effectiveness of remote sensing as a water management tool. However the accuracy of the universal algorithms presented in the literature was significantly lower than that of the empirical algorithm for Esthwaite Water. Universal algorithms therefore require further development. This study has shown that class type is an important factor in algorithm development. It is therefore likely that successful universal algorithms will require some consideration of the dominant class in the target bloom.

Application of the class identification routines requires forethought. It has already been suggested that *a priori* knowledge of the likely dominant class in a water body will make the task of confirming the identification of that class considerably easier. However, if this necessitated field visits, then this would mean that part of the benefit of organising a remote sensing campaign would be lost and operational costs would be increased. In Figure 1 the idea of successional phytoplankton development over the period of a year was introduced (Reynolds 1980, 1988). Given information on the prevailing environmental conditions prior to the campaign (principally sunlight, temperature and wind which could be provided by telemetered automatic monitoring stations or inferred from weather forecasts) it may be possible to determine which functional groups of phytoplankton are likely to be in ascendance.

As such Figure 1 forms the basis for an expert system which could be used prior to the remote sensing campaign to determine which classes are likely to be present. By discounting classes that are unlikely to be present the accuracy of the classification routines could be improved without recourse to expensive fieldwork.

This has been primarily an empirical study of phytoplankton optical properties. A more sophisticated approach would be the development of reflectance models of phytoplankton reflectance. Such models exist (Gordon *et al.* 1975 and Harwar 1995). They require accurate measurements of the inherent optical properties of phytoplankton for input, especially of the individual pigment *in vivo* absorption spectra. Were this input data available these models could be used to address how changes in the concentration of class marker pigments affected the reflectance spectrum. The measurements made for this study would only partially meet this pre-requirement for a modelling based assessment of classification from remote sensing reflectance spectra.

The presence of phytoplankton pigments can not be categorically confirmed using optical techniques. HPLC would have allowed confirmation of the presence and concentration of the different phytoplankton accessory pigments present in the species considered within this study (Mantoura and Llewellyn 1983). This would have improved the development of the identification routines. However, although HPLC is effective at identifying the presence of certain groups of pigments there are many accessory pigments which can not be separated using HPLC techniques. Furthermore HPLC analysis does not provide information on the absorption properties of the identified pigments and it is these properties which have led to the successful development of classification routines.

The reliance on published pigment absorption properties was a significant drawback to this study as these could not be categorically verified. Further investigation of the *in vivo* absorption properties of phytoplankton pigments is warranted. Given the molecular structure of the pigments it is possible to predict their absorption properties on physico-chemical grounds (M.H.Charlton, Zenica Industries, Manchester, U.K. pers. comm.). Therefore, following HPLC analysis to determine the pigment composition, a composite $a_{PHY}(\lambda)$ could be generated and used to construct a library of pure culture $a_{PHY}(\lambda)$. This could be used to develop and test more rigorous identification routines.

This study has concentration upon the measurement of pure culture $R_{PHY}(\lambda)$ in the experimental tank. As the work on Esthwaite Water showed, natural lakes tend to sustain a variety of phytoplankton classes with each class assuming dominance as seasonal chemical and physical changes provide the optimum growing conditions for that class. It is therefore more likely that the reflectance spectra from natural waters will be a composite of the reflectance spectra from a range of species drawn from various classes. To ensure that the classification routines could accommodate these composite reflectance spectra measurements of controlled mixtures of two or more classes could be made in the

experimental tank. These spectra could then be used to develop and test the classification routines. Similarly, both the experimental tank and Esthwaite Water were Case I waters. The classification and quantification routines need to be able to cope with Case II waters if they are to be universally applicable.

This study has concentrated upon the identification of phytoplankton class. There are more freshwater phytoplankton classes than those considered during this study and the 4^{th} derivative classification technique requires testing on a more diverse range of classes. However, because the technique was based upon pigment compositions differing between classes, additional classes with their own marker pigment features should be easily accommodated by the technique. The ability to remotely sense to the species level would provide even more useful information to the water manager. To achieve this a greater knowledge of the specific optical properties and pigment compositions of the freshwater phytoplankton species is clearly required.

Appendix One

The pigments reported as occurring in the five classes considered in this study.

Table 56. The pigments associated with Bacillariophytes.

Pigment	*Reference*
β-carotene	Rowan (1989)
CHL*a*	Rowan (1989)
CHLc_1	Rowan (1989)
CHLc_2	Rowan (1989)
CHLc_3	Rowan (1989)
DDX	Goodwin (1980)
DTX	Goodwin (1980)
FCX	Goodwin (1980), Claustre (1994)
VLX	Owens *et al.* (1987)
Zeaxanthin	Rowan (1989)

Table 57. The pigments associated with Chlorophytes.

Pigment	*Reference*
α-CART	Rowan (1989)
ATX	Rowan (1989)
ASX	Goodwin (1980), Britton (1983)
CTX	Goodwin (1980)
CHL*a*	Rowan (1989)
CHL*b*	Rowan (1989)
LUT	Rowan (1989)
SPX	Rowan (1989)
SPN	Rowan (1989)
TAX	Rowan (1989)
VLX	Owens *et al.* (1987)

Table 58. The pigments associated with Cryptophytes.

Pigment	*Reference*
α-CART	Rowan (1989), Goodwin (1980)
β-CART	Rowan (1989)
APC	Rowan (1989)
ALX	Rowan (1989), Claustre (1994)
ADX	Rowan (1989)
CHL*a*	Rowan (1989)
CHLc_2	Rowan (1989)
CHL*d*	Rowan (1989), Lee (1989), Britton (1983)
FCX	Rowan (1989)
LUT	Rowan (1989), Goodwin (1980)
PCC	Rowan (1989)
PCE	Rowan (1989)
ZEX	Rowan (1989), Goodwin (1980)

Table 59. The pigments associated with Cyanophytes.

Pigment	*Reference*
β-CART	Rowan (1989)
APC	Rowan (1989), Lee (1989)
APN	Healey (1968)
β-CTX	Goodwin (1980)
CLX	Rowan (1989)
CTX	Rowan (1989), Healey (1968)
CHL*a*	Rowan (1989)
C-PCC	Lee (1989)
C-PCE	Lee (1989)
ECH	Lee (1989), Rowan (1989), Goodwin (1980), Healey (1968)
LUT	Healey (1968)
MYX	Healey (1968)
MXX	Lee (1989), Rowan (1989), Goodwin (1980), Healey (1968)
PCC	Rowan (1989)
PCE	Rowan (1989)
PHC	Lee (1989), Rowan (1989)
ZEX	Rowan (1989), Goodwin (1980), Healey (1968) Claustre (1994)

Table 60. The pigments associated with Euglenophytes.

Pigment	*Reference*
CHL*a*	Rowan (1989)
CHL*b*	Rowan (1989)
DDX	Rowan (1989)
DTX	Rowan (1989)
LUT	Rowan (1989)
SIP	Rowan (1989)

Appendix Two

Locations of centres for pigment absorption features taken from the literature.

Table 61. The pigment database compiled during a review of the literature. The $a_{PIG}(\lambda)$ properties listed within this database are taken verbatim from the cited reference, there is no means of comprehensively verifying all the features listed, consequently the precision and accuracy of these features must be treated with caution.

nm	*Pigment*	*Abbreviation*	*Source*
412	Chlorophyll-a	CHLa	Hoepffner and Sathyendranath (1991)
418	Chlorophyll-a	CHLa	Goedheer (1970)
419	Chlorophyll-a	CHLa	Prezelin (1980)
435	Chlorophyll-a	CHLa	Hoepffner and Sathyendranath (1991)
437	Chlorophyll-a	CHLa	Prezelin and Boczar (1986)
437	Chlorophyll-a	CHLa	Goedheer (1970)
438	Chlorophyll-a	CHLa	Prezelin and Alberte (1978)
439	Chlorophyll-a	CHLa	Smith and Alberte (1994)
440	Chlorophyll-a	CHLa	Yentsch and Yentsch (1984)
440	Chlorophyll-a	CHLa	Bricaud et al. (1983)
440	Chlorophyll-a	CHLa	Owens et al. (1987)
440	Chlorophyll-a	CHLa	Johnsen et al. (1994)
440	Chlorophyll-a	CHLa	Bidigare et al. (1989)
440	Chlorophyll-a	CHLa	Davis-Colley et al. (1986)
440	Chlorophyll-c	CHLc	Davies-Colley et al. (1986)
454	Chlorophyll-c_2	CHLc	Prezlin and Boczar (1986)
460	Chlorophyll-c	CHLc	Mann and Myers (1968)
460	Diadinoxanthin	DDX	Johnsen et al. (1994)
465	Chlorophyll-c	CHLc	Owens et al. (1987)
465	Chlorophyll-c	CHLc	Johnsen et al. (1994)
466	β-carotene	β-CART	Smith and Alberte (1994)
467	Chlorophyll-c	CHLc	Bidigare et al. (1989)
470	Chlorophyll-b	CHLb	Johnsen et al. (1994)
470	Chlorophyll-b	CHLb	Kan and Thornber (1976)
470	Chlorophyll-b	CHLb	Owens et al. (1987)
470	Peridinin	PRD	Faust and Norris (1982)
474	Chlorophyll-b	CHLb	Smith and Alberte (1994)
475	Carotenoids	CART	Bricaud et al. (1983)
483	Chlorophyll-b	CHLb	Bidigare et al. (1989)
487	Theraxanthin	TRX	Smith and Alberte (1994)
490	Carotenoids	CART	Owens et al. (1987)
490	Diadinoxanthin	DDX	Johnsen et al. (1994)
490	Luteoxanthin	LTX	Owens et al. (1987)
490	Violaxanthin	VLX	Smith and Alberte (1994)
490	Violaxanthin	VLX	Owens et al. (1987)
492	Phycourobilin	PUB	Bidigare et al. (1989)
495	Phycourobilin	PUB	Bidigare et al. (1989)
495	R-Phycoerythrin	PCE	Kirk (1994)
496	Phycourobilin	PUB	Morel et al. (1993)
498	B-Phycoerythrin	PCE	Kirk (1994)
498	R-Phycoerythrin	PCE	Prezlin and Boczar (1986)
498	R-Phycoerythrin	PCE	Kirk and Tilney-Bassett (1978)
500	Fucoxanthin	FCX	Davis-Colley et al. (1986)
503	Myxoxanthophyll	MYX	Richardson et al. (1994)
505	Phycoerythrin	PCE	Yentsch and Yentsch (1984)

nm	*Pigment*	*Abbreviation*	*Source*
505	Phycoerythrin	PCE	Yentsch and Yentsch (1984)
525	Fucoxanthin	FCX	Bidigare et al. (1989)
526	Fucoxanthin	FCX	Bidigare et al. (1989)
530	Fucoxanthin	FCX	Hoeffpner and Sathyendranath (1991)
535	Fucoxanthin	FCX	Johnsen et al. (1994)
540	Fucoxanthin	FCX	Smith and Alberte (1994)
540	R-Phycoerythrin	PCE	Prezlin and Boczar (1986)
540	R-Phycoerythrin	PCE	Kirk and Tilney-Bassett (1978)
540	R-Phycoerythrin	PCE	Kirk (1994)
540	Siphonaxanthin	SPX	Kirk (1994)
540	Siphonaxanthin	SPX	Smith and Alberte (1994)
543	Phycoerythrin	PCE	Bidigare et al. (1989)
544	Fucoxanthin	FCX	Yentsch and Yentsch (1984)
544	Peridinin	PRD	Yentsch and Yentsch (1984)
544	Phycoerythrin	PCE	Kirk (1994)
544	R-Phycoerythrin	PCE	Prezlin and Boczar (1986)
545	B-Phycoerythrin	PCE	Kirk (1994)
545	B-phycoerythrin	PCE	Prezlin and Boczar (1986)
545	Fucoxanthin	FCX	Goedheer (1976)
545	Peridinin	PRD	Hoepffner and Sathyendranath (1991)
546	B-Phycoerythrin	PCE	Kirk and Tilney-Bassett (1978)
546	B-Phycoerythrin	PCE	Kirk (1994)
548	Phycoerythrobilin	PEB	Bidigare et al. (1989)
550	Phycoerythrobilin	PEB	Morel et al. (1993)
552	R-Phycoerythrin	PCE	Kirk (1994)
553	R-Phycocyanin	PCC	Prezlin and Boczar (1986)
553	R-Phycocyanin	PCC	Kirk (1994)
553	R-Phycocyarin	PCC	Kirk and Tilney-Bassett (1978)
555	Phycoerythrin	PCE	Kirk (1994)
555	R-Phycoerythrin	PCE	Prezlin and Boczar (1986)
563	B-Phycoerythrin	PCE	Kirk (1994)
563	B-phycoerythrin	PCE	Prezlin and Boczar (1986)
563	C-Phycoerythrin	PCC	Kirk (1994)
565	C-Phycoerythin	PCC	Prezlin and Boczar (1986)
565	C-Phycoerythrin	PCE	Lee (1989)
565	Phycoerythrin	PCE	Kirk (1994)
565	Phycoerythrin	PCE	Kirk (1994)
565	R-Phycoerythrin	PCE	Prezlin and Boczar (1986)
568	Phycoerythrin	PCE	Kirk (1994)
568	Phycoerythrocyanin	PCEC	Lee (1989)
568	Phycoerythrocyanin	PCEC	Kirk (1994)
568	R-Phycoerythrin	PCE	Prezlin and Boczar (1986)
568	R-Phycoerythrin	PCE	Kirk (1994)
570	Phycoerythrin	PCE	Faust and Norris (1982)
575	Phycoerythrin	PCE	Yentsch and Yentsch (1984)
580	Phycoerythrin	PCE	Bricaud et al. (1988)
585	Phycocyanin	PCC	Kirk (1994)
585	Phycocyanin	PCC	Kirk (1994)
585	Phycocyanin	PCC	Kirk (1994)
586	Chlorophyll-c	CHLc	Johnsen et al. (1994)
586	Chlorophyll-c_1	CHLc	Johnsen et al. (1994)
586	Chlorophyll-c_2	CHLc	Johnsen et al. (1994)
586	Chlorophyll-c_3	CHLc	Johnsen et al. (1994)
590	Carotenoids	CART	Maaske and Haardt (1987)

nm	*Pigment*	*Abbreviation*	*Source*
590	Phycoerythrocyanin	PCEC	Kirk (1994)
598	R-Phycoerythrin	PCE	Kirk (1994)
615	Phycocyanin	PCC	Yentsch and Yentsch (1984)
615	Phycocyanin	PCC	Yentsch and Yentsch (1984)
615	Phycocyanin	PCC	Kirk (1994)
615	R-Phycocyanin	PCC	Prezlin and Boczar (1986)
615	R-Phycocyanins	PCC	Prezlin and Boczar (1986)
617	C-Phycocyanin	PCC	Prezlin and Boczar (1986)
618	B-Allophycocyanin	APC	Kirk (1983)
618	B-Allophycocyanin	APC	Kirk (1994)
618	Chlorophyll-a	CHLa	Prezelin (1980)
618	Chlorophyll-a	CHLa	Goedheer (1976)
618	R-Phycocyanin	PCC	Kirk (1994)
620	Allophycocyanin	APC	Kirk (1994)
620	C-Phycocyanin	PCC	Kirk (1994)
620	C-Phycocyanin	PCC	Lee (1989)
620	Chlorophyll-c	CHLc	Owens et al. (1987)
620	Phycocyanin	PCC	Bricaud et al. (1988)
620	Phycocyanin	PCC	Kirk (1994)
620	Phycocyanin	PCC	Kirk (1994)
623	Chlorophyll-a	CHLa	Hoepffner and Sathyendranath (1991)
625	Chlorophyll-a	CHLa	Owens et al. (1987)
625	Phycocyanin	PCC	Faust and Norris (1982)
627	Phycobilin	PCB	Davis-Colley et al. (1986)
630	Chlorophyll-c	CHLc	Bidigare et al. (1989)
630	Chlorophyll-c	CHLc	Davis-Colley et al. (1986)
630	Phycocyanin	PCC	Kirk (1994)
630	R-Phycocyanins	PCC	Prezlin and Boczar (1986)
632	Chlorophyll-c	CHLc	Bricaud et al. (1983)
635	Chlorophyll-c	CHLc	Smith and Alberte (1994)
635	Chlorophyll-c	CHLc	Johnsen et al. (1994)
640	Chlorophyll-b	CHLb	Prezlin and Boczar (1986)
640	Chlorophyll-b	CHLb	Kirk and Tilney-Basset (1978)
640	Chlorophyll-c	CHLc	Mann and Myers (1968)
643	Chlorophyll-b	CHLb	Burger-Weirsma et al. (1986)
645	Phycocyanin	PCC	Kirk (1994)
645	R-Phycocyanins	PCC	Prezlin and Boczar (1986)
648	Chlorophyll-b	CHLb	Prezlin and Boczar (1986)
649	Chlorophyll-b	CHLb	Johnsen et al. (1994)
650	Allophycocyanin	APC	Lee (1989)
650	Allophycocyanin	APC	Kirk (1994)
650	Allophycocyanin	APC	Kirk (1994)
650	Allophycocyanin	APC	Prezlin and Boczar (1986)
650	Chlorophyll-b	CHLb	Johnsen et al. (1994)
650	Chlorophyll-b	CHLb	Bidigare et al. (1989)
650	Chlorophyll-b	CHLb	Bricaud et al. (1983)
650	Chlorophyll-b	CHLb	Owens et al. (1987)
650	Chlorophyll-b	CHLb	Kirk and Tilney-Basset (1978)
651	Chlorophyll-b	CHLb	Davis-Colley et al. (1986)
652	Chlorophyll-b	CHLb	Kan and Thornber (1976)
654	Allophycocyanin	APC	Prezlin and Boczar (1986)
656	B-Phycoerythrin	PCE	Kirk (1994)
658	Chlorophyll-b	CHLb	Smith and Alberte (1994)
662	Chlorophyll-a	CHLa	Prezlin and Boczar (1986)

nm	*Pigment*	*Abbreviation*	*Source*
662	Chlorophyll-a	CHLa	Kirk and Tilney-Basset (1978)
670	Chlorophyll-a	CHLa	Prezlin and Boczar (1986)
670	Chlorophyll-a	CHLa	Kirk and Tilney-Basset (1978)
670	Chlorophyll-a	CHLa	Lee (1989)
671	B-Allophycocyanin	APC	Kirk and Tilney-Bassett (1978)
671	B-Allophycocyanin	APC	Kirk (1994)
672	Chlorophyll-a	CHLa	Owens et al. (1987)
673	Chlorophyll-a	CHLa	Goedheer (1976)
675	Chlorophyll-a	CHLa	Faust and Norris (1982)
675	Chlorophyll-a	CHLa	Yentsch and Yentsch (1984)
675	Chlorophyll-a	CHLa	Prezelin and Alberte (1978)
675	Chlorophyll-a	CHLa	Hoepffner and Sathyendranath (1991)
675	Chlorophyll-a	CHLa	Prezelin (1980)
675	Chlorophyll-a	CHLa	Johnsen et al. (1994)
675	Chlorophyll-a	CHLa	Bidigare et al. (1989)
675	Chlorophyll-a	CHLa	Bricaud et al. (1983)
676	Chlorophyll-a	CHLa	Davis-Colley et al. (1986)
677	Chlorophyll-a	CHLa	Kirk and Tilney-Basset (1978)
678	Chlorophyll-a	CHLa	Prezlin and Boczar (1986)
680	Chlorophyll-a	CHLa	Lee (1989)
681	Chlorophyll-a	CHLa	Smith and Alberte (1994)
684	Chlorophyll-a	CHLa	Prezlin and Boczar (1986)
684	Chlorophyll-a	CHLa	Kirk and Tilney-Basset (1978)
692	Chlorophyll-a	CHLa	Kirk and Tilney-Basset (1978)
705	Chlorophyll-a	CHLa	Kirk and Tilney-Basset (1978)

Appendix Three

The formulations for the two culturing media used in this study.

The ingredients were added, at the specified concentrations to spherical glass flasks containing sand-filtered de-ionised water and made up to 1 to 3 L depending upon the volume of culture required. It was necessary to adjust the pH of the DM media to 6.9 using drops of 1M HCl. The flasks were then sealed with a sponge bung and wrapped in aluminium foil and autoclaved at a minimum of 15 psi for a minimum of 15 minutes.

Table 62. Jarworski's Blue-green Medium (JM) Final solution: 1 mL of each stock added to 1 L deionized water.

Stock solution		*g/200 mL deionized water*
1	$Ca(NO_3)_2.4H_2O$	4.0
2	KH_2PO_4	2.48
3	$MgSO_4.7\ H_2O$	10.0
4	$NaHCO_3$	3.18
5	EDTA FeNa	0.45
	EDTA Na_2	0.45
6	H_3BO_3	0.496
	$MnCl_2.2\ H_2O$	0.278
	$(NH_4)_6Mn_7O_{24}.4\ H_2O$	0.2
7	Cyanocabalanin B_{12}	0.008
	Thiamin B_1	0.008
	Biotin	0.008
8	$NaNO_3$	16.0
9	$Na_2HPO_4.12\ H_2O$	7.2

Table 63. Diatom Media (DM). Final solution: 1 mL of each stock solution added to 992 mL of deionized water.

Stock solution		*g/200 mL deionized water*
1	$Ca(NO_3)_2.4\ H_2O$	4.0
2	KH_2PO_4	2.48
3	$MgHCO_3.7\ H_2O$	5.0
4	$NaHCO_3$	3.18
5	EDTA FeNa	0.45
	EDTA Na_2	0.45
6	H_3BO_3	0.496
	$MnCl_2.4\ H_2O$	0.278
	$(NH_4)_6Mn_7O_{24}.4\ H_2O$	0.2
7	Cyanocabalanin B_{12}	0.008
	Thiamin B_1	0.008
	Biotin	0.008
8	$NaSiO_3.9H_2O$	11.4

Bibliography

Aas E 1981. The refractive index of phytoplankton. Inst. Rep. Ser. Univ. Oslo **46**, 61-81.

Agusti S 1991. Allometric scaling of light absorption and scattering by phytoplankton cells. *Canadian Journal Fish Aquatic Science*, **48**(5), 763-767.

Ahn Y-H, Bricaud A, Morel A 1992. Light backscattering efficiency and related properties of some phytoplankters. *Deep-Sea Research*, **39**(11/12), 1835-1855.

Almanza E, Melack JM 1985. Chlorophyll differences in Mono Lake (California) observable on Landsat imagery. *Hydrobiologia*, **122,**13-17.

Andre J 1992. Ocean colour remote-sensing and the subsurface vertical structure of phytoplankton pigments. *Deep-Sea Research*, **39**(5), 763-779.

Babin M, Morel A, Gentili B 1996. Remote sensing of sea surface sun-induced chlorophyll fluorescence: consequences of natural variations in the optical characteristics of phytoplankton and the quantum yield of chlorophyll-a fluorescence. *International Journal of Remote Sensing*, **17**(12), 2417-2448.

Baker KS, Smith RC 1990. Irradiance transmittence through the air-water interface. *Ocean Optics X SPIE*, **1302**, 556-565.

Bannister TT 1988. Estimation of absorption coefficients of scattering suspensions using opal glass. *Limnology and Oceanography*, **33**(4 part 1), 607-615.

Beakes GW, Canter HM, Jarworski GHM 1988. Zoospore ultrastructure of *Zygorhizidium affluens* and *Zygorhizidium planktonicum*, two chytrids parasitizing the diatom *Asterionella formosa*. *Canadian Journal of Botany*, 66(6), 1054-1067.

Bidigare RR, Morrow JH, Kiefer DA 1989. Derivative analysis of spectral absorption by photosynthetic pigments in Western Sargasso Sea. *Journal Marine Research*, 47, 323-341.

Boddy L, Morris C, Wilkins M, Tarran G, Burkill P 1994. Neural network analysis of flow cytometric data for 40 marine phytoplankton species. *Cytometry*, 15, 283-293.

Bold HC, Wynne MJ 1985. *Introduction to the algae. Structure and reproduction.* 2nd edition. New Jersey. Prentice and Hall.

Brown CW, Esaias WE, Thompson AM 1995. Predicting phytoplankton composition from space - using the ratio of euphotic depth to mixed-layer depth: an evaluation. *Remote Sensing Environment*, **53**, 172-176.

Bricaud A, Morel A, Prieur L 1981. Absorption by dissolved organic matter of the sea (yellow substance) in the UV and visible domains. *Limnology and Oceanography*, **26**(1), 43-53.

Bricaud A, Morel A, Prieur L 1983. Optical efficiency factors of some phytoplankters. *Limnology and Oceanography*, **28**(5), 816-832.

Bricaud A, Morel A 1986. Light attenuation and scattering by phytoplanktonic cells: a theoretical model. *Applied Optics*, **25**(4), 571-580.

Bricaud A, Bedhomme AL, Morel A 1988. Optical properties of diverse phytoplanktonic species: experimental results and theoretical interpretation. *Journal of Plankton Research*, **10**(5), 851-873.

Bricaud A, Stramski D 1990. Spectral absorption coefficients of living phytoplankton and non algal biogenous matter: a comparison between the Peru upwelling area and the Sargasso Sea. *Limnology and Oceanography*, **35**(3), 562-582.

Britton G 1989. *The biochemistry of natural pigments*. Cambridge University Press.

Bukata RP, Jerome JH, Bruton JE, Jain SC, Zwick HH 1981. Optical water quality model of Lake Ontario. 1: Determination of the optical cross sections of organic and inorganic particulates in Lake Ontario. *Applied Optics*, **20**(9), 1696-703.

Bukata RP, Bruton JE, Jerome JH 1985. Application of direct measurements of optical parameters to the estimation of lake water quality indicators. *Scientific series No. 140 Inland waters directorate National water research Institute of Canada.*

Burger-Wiersma T, Post AF 1989. Functional analysis of the photosynthetic apparatus of *Prochlorothix hollandia* (*Prochlorales*) a chlorophyll-*b* containing Procaryote. *Plant Physiology*, **91,** 770-774.

Butterwick C, Heaney SI, Talling JF 1982. A comparison of eight methods for estimating the biomass and growth of planktonic algae. *Br. Phycol. J.*, **17**, 69-79.

Butler WL, Hopkins DW 1970. Higher derivative analysis of complex absorption spectra. *Photochemistry and Photobiology*, **12**, 439-450

Carder KL, Steward RG, Paul JH, Vargo GA 1986. Relationships between chlorophyll and ocean color constituents as they affect remote-sensing reflectance models. *Limnology and Oceanography*, **31**(2), 403-413.

Carder KL, Steward RG, Harvey GR, Ortner PB 1989. Marine humic and fulvic acids: their effects on remote sensing of ocean chlorophyll. *Limnology and Oceanography*, **34**(2), 68-81.

Charlton FL 1990. *A radiometric assessment of water quality in the English Lake District.* Unpublished MSc dissertation, Sutton Bonnington, Nottingham University.

Christiansen T 1962. The gross classification of algae. *Algae and Man.* Ed, Jackson DF. Plenum Press New York, 59-64.

Claustre H 1994. The trophic status of various oceanic provinces as revealed by phytoplankton pigment signatures. *Limnology and Oceanography*, **39**(5), 1206-1210.

Cooley WW, Lohnes PR 1971. *Multivariate data analysis.* J Wiley and Sons.

Curran PJ 1994. Imaging spectrometry. *Progress in Physical Geography*, **18**(2), 247-266.

Davis-Colley RJ, Pirdmore RD, Hewitt JE 1986. Optical properties of some freshwater phytoplanktonic algae. *Hydrobiologia*, **133**, 165-178.

Davis-Colley RJ, Vant WN 1987. Absorption of light by yellow substance in freshwater lakes. *Limnology and Oceanography*, **32**(2), 416-425.

Dekker AG, Hoogenboom HJ, Volten H, Scheurs R, de Haan JF 1997. Angular scattering functions of algae and silt: an analysis of backscattering to scattering fraction. *Ocean Optics XIII*, **SPIE Vol. 2963**, 392-400.

Dekker AG, Malthus TJ, Wijnen MM, Seyhan E 1991. A multi-sensor approach to spectral signature analysis of inland waters. *Proceedings of the 5th International Colloquium - Physical measurements and signatures in remote sensing, Courcheval, France, 14-18 January 1991.* 177-180.

Dekker AG, Malthus TJ, Winen M, Seyhan E 1992a. Remote sensing as a tool for assessing water quality in Loosdrecht Lakes. *Hydrobiologia*, **233**, 137-159.

Dekker AG, Malthus TJ, Wijnen MM, Seyhan E 1992b. The effect of spectral bandwidth and positioning on the spectral signature analysis of inland waters. *Remote Sensing Environment*, **41**, 211-225.

Dekker AG, Malthus TJ, Wijen MM 1992c. Spectral band location for remote sensing of turbid and/or euthrophic waters. *Proceedings of the 1st thematic conference on remote sensing for marine and coastal environments: Needs and solutions for pollution monitoring, control and abatement. New Orleans, Louisiana, USA. 15-17 June 1992.* Volume II, SPIE Volume 1930, 955-970

Dekker AG 1993. *Detection of optical water quality parameters for eutrophic waters by high resolution remote sensing.* Unpublished PhD thesis. Vrije Universiteit, Amsterdam.

Dekker AG, Peters SWM 1993. The use of the thematic mapper for the analysis of eutrophic lakes: a case study in the Netherlands. *International Journal of Remote Sensing*, **14**(5), 799-821.

Demetriades-Shah TH, Steven MD, Clark JA 1990. High resolution derivative spectra in remote sensing. *Remote Sensing Environment*, **33**, 55-64.

Doucha J, Kubin S 1976. Measurement of *in vivo* absorption spectra of microscopic algae using bleached cells as a reference sample. *Arch. Hydrobiol. Suppl.*, **49**, 199-213.

Duysens LNM 1956. The flattening of the absorption spectrum of suspensions as compared to that of solutions. *Biochemica et Biophysica Acta*, **19**, 1-12.

Eberhart RC, Dobbins RW 1990. *Neural network PC tools: A practical guide.* Academic Press Inc.

Edmondson WT 1980. Secchi disk and chlorophyll. *Limnology and Oceanography*, **25**(2), 378-379.

Effler SW, Roop R, Perkins MG 1988. A simple technique for estimating absorption and scattering coefficients. *Water Resources Bulletin*, **24**(2), 397-404.

Effler SW, Perkins MG, Wagner BA 1991. Optics of Little Sodus Bay. *Journal Great Lakes Research*, **17**(1), 109-119.

Enke CG, Nieman TA 1976. Signal-to-noise ratio enhancement by least-squares polynomial smoothing. *Analytical Chemistry*, **48**(8), 705-712.

Faust MA, Norris KH 1982. Rapid *in vivo* spectrophotometric analysis of chlorophyll pigments in intact phytoplankton cultures. *British Phycological Journal*, **17**, 351-361.

Foody GM, Arora MK 1995. An evaluation of the factors affecting neural network classification accuracy. *Proceedings of the 21st Annual Conference of the Remote Sensing Society, 11-14 September 1995, University of Southampton*. 42-49.

Foody GM 1996. Approaches for the production and evaluation of fuzzy land cover classification from remotely-sensed data. *International Journal of Remote Sensing*, **17**(7), 1317-1340.

French CS, Brown JS, Prager L, Lawerence M 1967. Analysis of spectra of natural chlorophyll complexes. *Carniegie Institution Washington Year B*, **67**, 536-546.

Ganf GG, Oliver RL, Walsby AE 1989. Optical properties of gas vacuolate cells and colonies of *Microcystis* in relation to light attenuation in a turbid stratified reservoir (Mount Bold Reservoir South Australia). *Australian Journal Marine Freshwater Research*, **40**, 595-611.

Ganf GG, Heany SI, Corry J 1991. Light absorption and pigment content in natural populations and cultures of a non-gas vacuolate cynanobacterium *Oscillatoria bourrellyi*. *Journal of Plankton Research*, **13**(5), 1101-1121.

George DG, Charlton FL 1996. *Inland use of airborne remote sensing*. NRA R&D Technical Report E5.

George G, Hewitt DP 1990. The remote sensing of phytoplankton in oliogtrophic lakes. *Proceedings of the NERC Symposium on Airborne Remote Sensing, British Geological Survey, Keyworth, Nottingham 18-19 December 1990*. 193-210.

George G 1993. Physical and chemical scales of pattern in freshwater lakes and reservoirs. *The Science of the Total Environment*, **135**, 1-15.

Geikes WWC, Kraay GW 1983. Dominance of Cryptophycea during the phytoplankton spring bloom in the central North Sea detected by HPLC analysis of pigments. *Marine Biology*, **75**, 179-185.

Gitelson A 1993. Algorithms for remote sensing of phytoplankton pigments in inland waters. *Advances in Space Research*, **13**(5), 197-201.

Gitelson A, Szilagyi F, Mittenzwey KH 1993. Improving quantitative remote sensing for monitoring of inland water quality. *Water Research*, **27**(7), 1185-1194.

Gitelson A, Mayo M, Yacobi YZ, Parparov A, Berman T 1994. The use of high-spectral resolution radiometer data for detection of low chlorophyll concentrations in Lake Kinneret. *Journal of Plankton Research*, **16**(8), 993-1002.

Goedheer JC 1976. Spectral properties of the Blue green alga *Anacystis nidulans* grown under different environmental conditions. *Photosynthetica*, **10**(4), 411-422.

Gons HJ, Kromkamp J, Rijkeboer M, Schofield O 1992. Characterisation of the light field in laboratory scale enclosures of eutrophic lake water (Lake Loosdrecht, The Netherlands). *Hydrobiologia*, **238**, 99-109.

Goodin DG, Han LH, Fraser EN 1993. Analysis of suspended solids in water using remotely sensed high-resolution derivative spectra. *Photogrammetric Engineering and Remote Sensing*, **59**(4), 505-510.

Goodwin TW 1980. *The biochemistry of the Carotenoids: Volume 1 Plants*. Chapman and Hall.

Gordon HG, Brown OB, Jacobs MM 1975. Computed relationships between the inherent and apparent optical properties of a flat homogenous ocean. *Applied Optics*, **14**(2), 417-427.

Gordon HR, MacCluny WR 1975. Estimation of the depth of sunlight penetration in the sea for remote sensing. *Applied Optics*, **14**, 413-416.

Gower JFR, Lin S, Borstad GA 1984. The information content of different optical spectral ranges for remote chlorophyll estimation in coastal waters. *International Journal of Remote Sensing*, **5(2)**, 340-364.

Gregg WW, Carder KL 1990. A simple spectral solar irradiance model for cloudless maritime atmospheres. *Limnology and Oceanography*, **8**(35), 1657-1675.

Haardt H, Maske H 1987. Specific *in vivo* absorption coefficient of chlorophyll-a at 675 nm. *Limnology and Oceanography*, **32**(3), 608-619.

Hakvoort JHM 1994. *Absorption of light by surface water*. Unpublish PhD thesis. Delft University of Technology, Delft, The Netherlands.

Hall GH, Parker JE, Deville MM, Lishman JP 1995. *Changes in the phosphorus content of the sediment of Esthwaite Water and water quality in response to a decreasing input of sewage-borne phosphorus*. Progress report to the NRA (North West Region).

Han L, Rundquist DC, Liu LL, Fraser RN 1994. The spectral response of algal chlorophyll in water with varying levels of suspended sediment. *International Journal of Remote Sensing*, **15**(18), 3707-3718.

Harding LW, Itsweire EC, Esaias WE 1992. Determination of phytoplankton chlorophyll concentrations in the Chesapeake Bay with aircraft remote sensing. *Remote Sensing Environment*, **40**, 79-100.

Harwar M 1995. *Proceedings of the 21st Annual Conference of the Remote Sensing Society, 11-14 September 1995, University of Southampton.*

Healey FP 1968. The carotenoids of four blue-green algae. *Journal of Phycology*, **4**, 126-129.

Heaney SI, Smyly WJP, Talling JF 1986. Interactions of physical chemical and biological processes in depth and time within a productive English lake during summer stratification. *International Revue der Gesamten Hydrobiologie*, **17**, 441-494.

Heaney SI, Corry JE, Lishman JP 1992. Changes of water quality and sediment phosphorous of a small productive lake following decreased phosphorus loading. *Eutrophication: research and application*. Ed. Sutcliffe DW and Jones JG. 119-131.

Hoepffner N, Sathyendranath S 1991. Effect of pigment composition on absorption properties of phytoplankton. *Marine Ecological Progress Series*, **73**, 11-23.

Hoepffner N, Sathyendranath S 1993. Determination of the major groups of phytoplankton pigments from the absorption spectra of total particulate matter. *Journal Geophysical Research*, **98** (c12), 22789-22803.

Hope K 1968. *Methods of multivariate analysis*. University London Press.

Hotelling 1935. The most predictable criterion. *Journal of educational psychology*, **26**, 139-142.

Hotelling 1936. Relations between two sets of variables. *Biometrika*, **28**, 321-377.

Hunt RWG 1992. *Measuring colour*. Ellis Horwood.

Jackson RD, Moran MS, Slater PN, Biggar SF 1987. Field calibration of reference reflectance panels. *Remote Sensing Environment*. **22**, 145-158.

Jakubauskas ME 1996. Canonical correlation analysis of coniferous forest spectral and biotic relations. *International Journal of Remote Sensing*, **17**(12), 2323-2332.

Jeffery SW, Humphrey GF 1975. New spectrophotometer equations for determing chlorophylls a, b, c1 and c2 in height plants algae and natural phytoplankton. *Biochem. Physiol. Pflanzen.*, **167**, 191-194.

Johnsen G, Samset O, Granskog L, Sakshaug E 1994. *In vivo* absorption characteristics in 10 classes of bloom-forming phytoplankton: taxonomic characteristics and responses to photoadaptation by means of discriminant and HPLC analysis. *Marine Ecology Progress* Series, **105**, 149-157.

Jones JG 1979. A guide to methods for estimating microbial numbers and biomass in freshwater. *FBA scientific publication* **39**.

Kan J and Thornber TD 1976. The light-harvesting chl-a/b protein complex of *Chlamydomonas reinhardii*. *Plant Physiology*, 57, 47-52.

Kendall M 1968. *A course in multivariate analysis*. London. Griffin.

Khan A 1987. Problems of smoothing and differentiation of data by least-squares procedures and possible solutions. *Analytical Chemistry*, **59**, 654-657.

Kirk JTO 1975. A theoretical analysis of the contribution of algal cells to the attenuation of light within natural waters. *New Phytologist*, **75**, 11-20.

Kirk JTO, Tilney-Bassett RAE 1978. *The Plastids: their chemistry structure growth and inheritance*. Elsevier Press.

Kirk JTO 1981a. Spectral absorption properties of natural waters: contribution of the soluble and particulate fractions to light absorption in some inland waters in SE Australia. *Australian Journal Marine Freshwater Research*, **31**, 287-296.

Kirk JTO 1981b. Monte carlo study of the nature of the underwater light field and the relationship between optical properties of turbid yellow waters. *Australian Journal Marine Freshwater Research*, **32**, 517-532.

Kirk JTO, Tyler PA 1986. The spectral absorption and scattering properties of dissolved and particulate components in relation to the underwater light field of some tropical Australian freshwaters. *Freshwater Biology*, **16**, 573-583.

Kirk JTO 1994. *Light and photosynthesis in the aquatic ecosystem*. Cambridge University Press.

Kishino M, Booth CR, Okami N 1984. Underwater radiant energy absorbed by phytoplankton detritus dissolved organic matter and pure water. *Limnology and Oceanography*, **29**(3), 340-349.

Kishino M, Takahashi M, Okami N, Ichimura S 1985. Estimation of the spectral absorption coefficients of phytoplankton in the sea. *Bulletin of Marine Science*, **37**(2), 634-42.

Kitchen JC, Zaneveld RV 1990. On the non-correlation of the vertical structure of light scattering and chlorophyll a in Case I waters. *Journal Geophysical Research*, **95**(11), 20237-20246.

Krijgsman J 1994. *Optical remote sensing of water quality parameters: Interpretation of reflectance spectra*. Unpublished thesis, Vrije Univ., Amsterdam.

Kuik F, Stammes P, Hovenier JW 1991. Experimental determination of scattering matrices of water droplets and quartz particles. *Applied Optics*, **30**, 4872-4881.

Lee RE 1989. *Phycology*. Cambridge University Press.

Lee Z, Carder KL, Hauves SK, Steward RG 1994. Model for the interpretation of hyperspectral remote-sensing reflectance. *Applied Optics*, **33**(24), 5721-5732.

Lewis HG, Cote S, Tatnall ARL 1995. Shape motion and contextual descriptors for a neural network cloud classifier. *Proceedings of the 21st Annual Conference of the Remote Sensing Society, 11-14 September 1995, University of Southampton*. 3-10.

Lund JWG 1959. A simple counting chamber for nanoplankton. *Limnology and Oceanography*, 4, 57-65.

Maaske H, Haardt H 1987. Quantitative in vivo absorption spectra of phytoplankton: Detrital absorption and comparison with fluorescence excitation spectra. *Limnology and Oceanography*, **32**(3), 620-633.

Malthus TJ, Dekker AG 1990. Spectral light attenuation in a hypertrophic lake system (Loosdrecht lakes the Netherlands). *Verh. Internat. Verein. Limnol.*, **24**, 711-714.

Malthus TJ, Madira AC, Steven MD, Andrieu B, Baret F 1995. Monitoring crop productivity in the presence of both crop stress and variation in background colour. *Proceedings International Colloquium Photosynthesis and Remote Sensing*, 28-30 August 1995, 223-230.

Malthus TJ, Place CJ, Bennett S, North S 1996. An evaluation of the Airborne Thematic Mapper Sensor for monitoring inland waters. *Proceedings 22nd Annual Conference of the Remote sensing society, Durham 11-14 September 1996*, 325-332.

Mann JE and Myers J 1968. On pigment, growth, and photosynthesis of *Phaedactylum tricornutum*. *Journal of Phycology*, **4**, 347-348.

Mantoura RFC, Llewellyn CA 1983. The rapid determination of algal chlorophyll and carotenoid pigments and their breakdown products in natural waters by reversed phase HPLC. *Analytica Chimica Acta*, **151**, 297-314.

Mantovani JE, Cabral AP 1992. Tank depth determination for water radiometric measurements. *International Journal of Remote Sensing*, **13**(14), 2727-2733.

Maracci G, Ooms M 1988. Optical properties of seawater bodies: measurements with an underwater radiometer and a high-resolution spectroradiometer. *Proceedings IGARSS 1988 Symposium, Edinburgh, Scotland, 13-16 September 1988 Ref. ESA SP-284 (IEEE 88CH2497-6)*, 1379-1380.

Martin AE 1957. Difference and derivative spectra. *Nature*, **4579**, 231-233.

Maske H, Haardt H 1987. Quantitative *in vivo* absorption spectra of phytoplankton: Detrital absorption and comparison with fluorescence excitation spectra. *Limnology and Oceanography*, **32**(3), 620-633.

Matthews AM, Boxall SR 1994. Novel algorithms for the determination of phytoplankton concentration and maturity. *Proceedings 2nd Thematic Conference on Remote Sensing*. 173-180.

Mie G 1908. Beitrage zur Optok truber Medien speziell kolloidaler metallosungen. *Ann. Physik (Leipzig)*, **25**, 377.

Millie DF, Ingram DA, Dionigi CP 1990. Pigment and photosynthetic responses of *O.Aagardhii* to photon flux density and spectral quality. *Journal Phycology*, **26**, 660-666.

Millie DF, Baker MC, Ticker CS, Vinyard BT, Dionigi CP 1992. High resolution airborne remote sensing of bloom-forming phytoplankton. *Journal Phycology*, **28**, 281-290.

Milton EJ 1987 Principles of field spectroscopy. *International Journal of Remote Sensing*, **8**(12), 1807-1827.

Monteith JL, Unsworth MH 1990. *Principles of environmental physics*. Edward Arnold.

Morel A, Prieur L 1977. Analysis of variations in ocean colour. *Limnology and Oceanography*, **22**(4), 709-723.

Morel A, Bricaud A 1981. Theoretical results concerning light absorption in a discrete medium and application to specific absorption of phytoplankton. *Deep-Sea Research*, **28**(11), 1375-1393.

Morel A, Bricaud A 1986. Inherent optical properties of algal cells including picoplankton: theoretical and experimental results. *Photosynthetic phytoplankton*. Ed. Platt T and Li WKW. *Canadian Bulletin Fishery and Aquatatic Sciences,* **214**, 521-559.

Morel A, Ahn Y, Partensky F, Vaulot D, Claustre H 1993. *Prochlorococcus* and *Synechococcus*: A comparative study of their optical properties in relation to their size and pigmentation. *Journal Marine Research*, **51**, 617-649.

Nelson JR, Robertson CY 1993. Detrital spectral absorption: laboratory studies of visible light effects on phytodetritus absorption bacterial spectral signal and comparison to field measurements. *Journal Marine Research*, **15**, 181-207.

Nelson NB, Prezelin BB 1990. Chromatic light effects and physiological modelling of absorption properties of *Heterocapsa pygmaea*. *Marine Ecology Progress Series*, **63**, 37-46.

Novo EMLU, Hansom JD, Curran PJ 1989. The effect of viewing geometry on the relationship between reflectance and suspended solids. *International Journal of Remote Sensing*, **10**(8), 1357-1372.

Novo EMLU, Steffan CA, Braga CZF 1991. Results of a laboratory experiment relating spectral reflectance to total suspended solids. *Remote Sensing Environment*, **36**, 67-72.

NRA Report 1990. *Toxic blue-green algae. A report to the National River Authority.*

NRA 1993 R & D Report 4. Boxall SR, Chaddock SE, Matthews A, Holden N. *Airborne remote sensing of coastal waters.*

O'Haver TC, Green GL 1976. Numerical error analysis of derivative spectrometry for the quantitative analysis of mixtures. *Analytical Chemistry*, **48**(2), 312-318.

O'Haver TC 1982. Derivative spectroscopy: Theoretical aspects. Plenary Lecture. *Derivative Spectroscopy*, **Annual proceedings 1982**, 22-45.

O'Haver TC 1991. An introduction to signal processing in chemical measurement. *Journal of Chemical Education*, **68**, a147-a150.

Owens TG, Gallagher JC, Alberte RS 1987. Photosynthetic light-harvesting function of violaxanthin in *Nannochloropsis Spp.* (*Eustigmatophyceae*). *Journal Phycology*, **25**, 79-85.

Paola JD, Schowengerdt RA 1995. A review and analysis of back propagation neural networks for classification of remotely-sensed multi-spectral imagery. *International Journal of Remote Sensing*, **16**(16), 3033-3058.

Pearsall WH, Hewitt T 1934. Light penetration into freshwater. *Journal of Experimental Biology*, **XI**(1), 89-93.

Petzold TL 1972. Volume scattering functions for selected ocean waters. University of California, San Diego, Scripps Inst Oceanogr. Visibility Lab. 72-78.

Porkess 1988. Dictionary of statistics.

Preisendorfer RW 1986. Secchi disk science: Visual optics of natural waters. *Limnology and Oceanography*, **31**(5), 909-926.

Prezelin BB and Alberte 1978. Photosynthetic characteristics and organisation of chlorophyll in marine dinoflagellates. *Proceedings of the National Academy of Science USA*, 75, 1801-1804.

Prezelin BB and Boczar BA 1986. Molecular bases of cell absorption and fluorescence in phytoplankton: potential applications to studies in optical oceanography. *Progress in Phycological Research*. (Ed. Round and Chapman), **4**(1), 349-465.

Price JC 1991. On the value of high spectral resolution measurements in the visible and near infrared. *Proceedings of the 5th International Colloquium - Physical measurements and signatures in remote sensing, Courcheval France 14-18 January 1991*, 131-136.

Prieur L, Sathyendranath S 1981. An optical classification of coastal and oceanic waters based on the specific spectral absorption curves of phytoplankton pigments dissolved organic matter and other particulate matter. *Limnology and Oceanography*, **26**(4), 671-689.

Quibell G 1991. The effect of suspensoids sediment on reflectance from freshwater algae. *International Journal of Remote Sensing*, **12**(1), 177-182.

Quibell G 1992. Estimating chlorophyll concentration using upwelling radiance from different freshwater algal genera. *International Journal of Remote Sensing*, **13**(4), 2611-2621.

Ramsey EW 1992. Remote sensing of water quality in active to inactive cooling water reservoirs. *International Journal of Remote Sensing*, **13**(18), 3465-6488.

Reckhow KH 1980. Techniques for exploring and presenting data applied to lake phosphorous concentration. *Canadian Journal Fishery and Aquatic Science*, **37**, 290-294.

Reynolds CS 1980. Phytoplankton assemblages and their periodicity in stratifying lake systems. *Holarctic Ecology*, **3**, 141-159.

Reynolds CS 1982. Phytoplankton periodicity: Its motivation mechanisms and manipulation. *Report Freshwater Biological Association*, **50**, 60-75.

Reynolds CS 1984. *The ecology of freshwater phytoplankton*. Cambridge University Press.

Reynolds CS 1988. The concept of ecological succession applied to seasonal periodicity of freshwater phytoplankton. *Verh. Internat. Verein. Limnol.*, **23**, 683-691.

Reynolds CS 1993. Scales of disturbance and their role in plankton ecology. *Hydrobiologia*, **249**, 157-171.

Reynolds CS, Hall GH, Parker JE, DeVille M, Cranwell P, Cubby P, Irish AE, James B, Lawlor AJ, Lishman JP, Maberly S 1995. *Observations on the water quality of the lakes of the Windermere Catchment 1994*. IFE report to the NRA (North West Region).

Reynolds CS, Hall GH, Parker JE, DeVille M, Cubby P, Irish AE, James B, Lawlor AJ, Maberly SC, Rigg E 1996. *The water quality of the lakes of the Windermere Catchment 1995*. IFE report to the NRA (North West Region).

Richardson LL, Liu CJ, Buisson D, Ambrosia V 1994. The detection of algal photosynthetic accessory pigments using airborne visible-infrared imaging spectrometer (AVIRIS) spectral data. *Proceedings 2nd Thematic conference on Remote Sensing for Marine and Coastal Environments, New Orleans, Louisiana, 31 January - 2 February 1994*, 161-172.

Richardson LL, Buisson D, Ambrosia V 1995. Use of remote sensing coupled with algal accessory pigment data to study phytoplankton bloom dynamics in Florida Bay. *Third Thematic Conference on Remote Sensing for Marine and Coastal Environments, Seattle, Washington, 18-20 September 1995*.

Richardson LL, Ambrosia VG 1997. Remote sensing of algal pigment to determine coastal phytoplankton dynamics in Florida Bay. *Fourth International Conference on Remote Sensing for Marine and Coastal Environments, Florida, 17-19 March 1997.*

Roesler CS, Perry MJ, Carder KL 1989. Modelling in situ phytoplankton absorption from total absorption spectra in productive inland marine waters. *Limnology and Oceanography*, **34**(8), 1510-1523.

Rowan KS 1989. *Photosynthetic pigments of algae*. Cambridge Press.

Rummelhart DE, Hinton GE, Williams RJ 1986. Learning internal representations by error propagation. In *Parallel Distributed Processing: Explorations in the Microstructure of cognition Volume 1: Foundations*. Ed. Rummelhart DE and McClelland J.

Rundquist DC, Schalles JF, Peake JS 1995. The response of volume reflectance to manipulated algal concentrations above bright and dark bottoms at various depths in an experimental pool. *Geocato International,* **10(4),** 5-14.

Sathyendranath S, Lazzara L, Prieur L 1987. Variations in the spectral values of specific absorption of phytoplankton. *Limnology and Oceanography*, **32**(2), 403-415.

Savitzky A, Golay M 1964. Smoothing and differentiation of data by simplified least square procedures. *Analytical Chemistry* **36**(8), 1627-1638.

Schalles JF, Scheibe FR, Starks PJ, Troeger WW 1997. Estimation of algal and suspended sediment loads (singly and combined) using hyperspectral sensors and integrated mesocosm experiments. *4th International Conference on Remote Sensing Marine and Coastal Environments, 18-19 March 1997, New Orleans*. 247-257.

Scheurs R 1996. *Light scattering by algae*. Unpublished undergraduate dissertation, Astronomy Department, Vrije Universiteit, Amsterdam.

Shapiro J 1993. Theory and practice in the control of algal blooms. *Urban Waterside Regeneration.* Ed. White KN, Bellinger EG, Saul AJ, Symes M, Hendry K, **38**, 350-356.

Shibata K 1958. Spectrophotometery of intact biological materials. *Journal Biochemistry*, **45**, 599-623.

Shukla SS, Rusling JF 1984. Analysing chemical data with computers: errors and pitfalls. *Analytical Chemistry*, **56**(12), 1347-1368.

Smith CM, Alberte RS 1994. Characterisation of in vivo absorption features of Chlorophyte, Phaeophyte and Rhodophyte algal species. *Marine Biology*, **118**, 511-521.

Smith RC, Baker K 1981. Optical properties of the clearest natural water (200 - 800 nm). *Applied Optics*, **20**, 177-184.

Smith MJ, Moore T, Dumville M 1995. The performance of neural nets in image rectification *Proceedings of the 21st Annual Conference of the Remote Sensing Society, University of Southampton, 11-14 September 1995*, 11-18.

Standard Methods 1992. *Standard methods for the examination of water and waste water*. Ed. Franson MAH. American Public Health Association, 1015, 15th Street, NW 20005, Washington DC.

Stephanidis CN, Cracknell AP, Hayes LWB, Sloggett DR 1995. The implementation of a self-organised neural network for algal bloom detection. *Proceedings of the 21st Annual Conference of the Remote Sensing Society, University of Southampton, 11-14 September 1995*, 19-25.

Talling J, Driver D 1963. Some problems in the estimation of chlorophyll-a in phytoplankton. *Proceedings of the Conference on Primary Productivity Measurement, Marine and Freshwater, Hawaii (1961) U.S. Atomic Energy Commission.* TID-7633, 142-146.

Talling J 1971. The underwater light climate as a controlling factor in the production of freshwater phytoplankton. *Mitt. International Ver. Limnol.*, **19**, 214-243.

Tebbutt THY 1983. *Principles of water quality control*. Pergamon Press.

Tilman D, Kilham SS, Kilman P 1982. Phytoplankton community ecology: the role of limiting nutrients. *Ann. Rev. Ecol. System*, **13**, 349-372.

Tipping E, Hilton J, James B 1988. Dissolved organic matter in Cumbrian lakes and streams. *Freshwater Biology*, **19**, 371-378.

Topliss BJ, Platt T 1986. Passive fluorescence and photosynthesis in the ocean: implications for remote sensing. *Deep-Sea Research*, **33**(7), 849-864.

Turrell G 1982. Coefficients for polynomial least-squares interpolation. *Applied Spectroscopy*, **36**(1), 62-67.

Van-de-Hulst HC 1956. *Light scattering by small* particles. Wiley and Sons, New York.

Van Stokkom, HTC Stokman, GNM Hovenier JW 1993. Quantitative use of passive optical remote sensing over coastal and inland water bodies. *International Journal of Remote Sensing*, **14**(3), 541-563.

Vesk M, Jeffrey SW 1977. Effect of blue-green light on photosynthetic pigments and chloroplast structure in unicellular marine algae from six classes. *Journal Phycology*, **13**, 280-288.

Volten H, de Haan JF, Hovenier JW, Scheurs R, Vassen W, Dekker AG, Hoogenboom HJ, Charlton FL, Wouts R 1997. Laboratory measurements of angular distributions of light scattered by phytoplankton and silt. *Submitted to Limnology and Oceanography*.

Whitlock CH, Kuo CY, LeCroy SR 1982. Criteria for the use of regression analysis for remote sensing of sediments and pollutants. *Remote Sensing Environment*, **12**, 151-168.

Wilhelm C, Rudolph I, Renner W 1991. A quantitative method based upon HPLC aided pigment analysis to monitor structure and dynamics of the phytoplankton assemblage - A study from Lake Meerfelder Maar. *Arch. Hydrobiol.*, **123(1)**, 21-35.

Wilson AK, Mockeridge W, White S 1996. The NERC airborne remote sensing facility: Development and implementation of the integrated data system. *Proceedings 22nd Annual Conference of the Remote sensing society, University of Durham, 11-14 September 1996.*

Wróbel D 1977. Second and fourth derivative absorption spectra of chlorophyll a and chlorophyll c in vivo and in vitro. *Photosynthetica*, **11**(1), 90-92.

Yentsch CM, Yentsch CS 1984. Emergence of optical instrumentation for measuring biological properties. *Oceanogr. Mar. Biol. Ann. Rev.*, **22**, 55-98.

Yentsch CS 1962. Measurement of visible light absorption by particulate matter in the ocean. *Limnology and Oceanography*, **7**, 207-217.

Youngman RE 1971. Algal monitoring of water supply reservoirs and rivers. *Tech. Mem. Wat. Res. Ass.*, **TM63**, 1-26.

Zacharias M, Neimann O, Borstad G 1992. An assessment and classification of a multispectral bandset for the remote sensing of intertidal seaweeds. *Canadian Journal Remote Sensing*, **18**, 263-274.